

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/EP05/002556

International filing date: 10 March 2005 (10.03.2005)

Document type: Certified copy of priority document

Document details: Country/Office: DE
Number: 10 2004 013 842.7
Filing date: 20 March 2004 (20.03.2004)

Date of receipt at the International Bureau: 11 April 2005 (11.04.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

BUNDESREPUBLIK DEUTSCHLAND

10. 03. 2005

**Prioritätsbescheinigung über die Einreichung
einer Patentanmeldung****Aktenzeichen:** 10 2004 013 842.7**Anmeldetag:** 20. März 2004**Anmelder/Inhaber:** Degussa AG, 40474 Düsseldorf/DE**Bezeichnung:** Nitrilhydratasen aus Metagenombibliotheken**IPC:** C 12 N 9/88

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 23. Februar 2005
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag

A handwritten signature in black ink, which appears to read "Werner". Below the signature, the name "Werner" is printed in a smaller, standard font.

Nitrilhydratasen aus Metagenombibliotheken

Die vorliegende Erfindung ist auf spezielle degenerierte Primer gerichtet. Diese werden vorzugsweise in einem Verfahren zur Herstellung von Nitrilhydratasen eingesetzt.

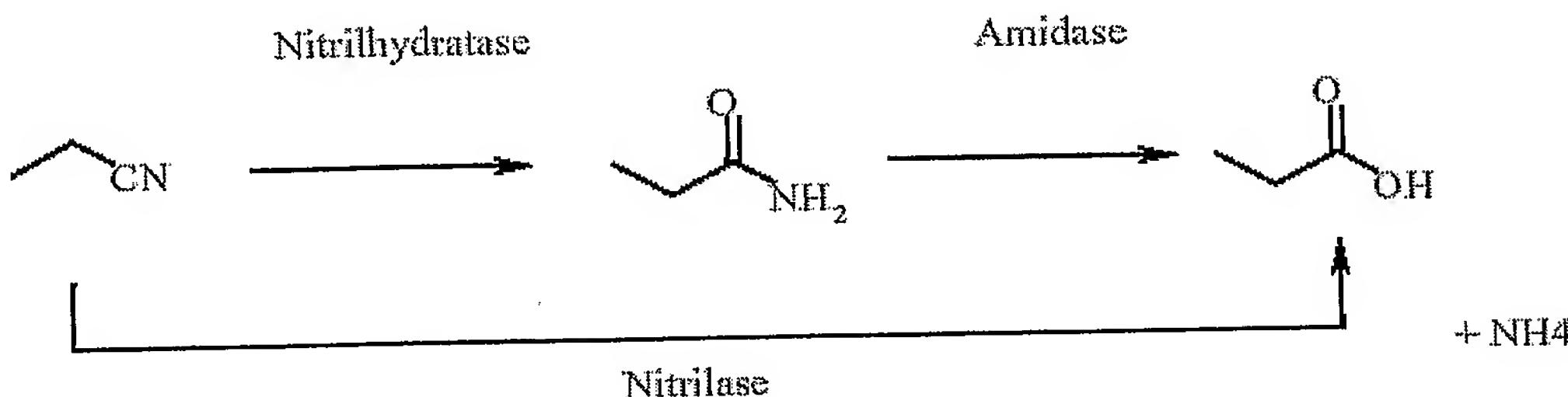
5 Gegenstand der vorliegenden Erfindung sind daher auch die durch das unter Zuhilfenahme der Primer durchgeführte Verfahren hergestellten Nitrilhydratasen und weitere für deren Aktivität benötigte Proteine. Ebenso bilden die diese Proteinsequenzen codierenden Nukleinsäuren und diese aufweisenden Expressionssysteme einen weiteren Gegenstand der Anmeldung. Die Verwendung der Nitrilhydratasen und der 10 zugrundeliegenden Nukleinsäuresequenzen bilden einen weiteren Aspekt der vorliegenden Erfindung.

Die Strukturklassen der Amide und Carbonsäuren gewinnen 15 mehr und mehr an Bedeutung als Vorstufen von Feinchemikalien. Spezielle Aminoamide und (proteinogene und nicht-proteinogene) Aminosäuren sind Schlüsselintermediate für die Synthese von pharmazeutischen und agrochemischen Produkten, als auch im Lebensmittelbereich. Insbesondere 20 enantiomerenreine Amide und Aminosäuren spielen eine immer größer werdende Rolle in den oben genannten Anwendungsbereichen.

Aminonitril-Vorstufen, wie sie für die Herstellung der oben angegebenen Verbindungsklassen benötigt werden, sind in 25 racemischer Form leicht über die so genannte Streckersynthese zugänglich. Die so gewonnenen Nitrile können anschließend mittels chemischer oder enzymatischer Verseifung in die entsprechenden Amide und Carbonsäuren überführt werden.

30 Es sind drei Enzyme bekannt, die an der enzymatischen Hydrolyse von Nitrilen beteiligt sein können. Nitrilasen setzen eine Nitril-Funktion direkt zur Säure um, wohingegen Nitrilhydratasen (E.C. 4.2.1.84) hier das entsprechende Amid bilden. Dieses kann durch eine Amidase (E.C. 3.5.1.4)

abschließend in die entsprechende Carbonsäure umgesetzt werden (Schema 1).



Schema 1:

5 Die Verseifung von Nitrilen zu den entsprechenden Amiden und Säuren mittels isolierter Enzyme oder Ganz-Zell-Katalysatoren hilft große Mengen Salz zu sparen, welche ansonsten bei dem Neutralisierungsschritt nach der chemischen Verseifung von Nitrilen anfallen würden. Aus 10 diesem Grund stellt die enzymatische Verseifung von Nitrilen zu z.B. Aminoamiden und/oder Aminosäuren ein nachhaltigeres Produktionsverfahren dar.

Nitrilhydratasen bestehen in ihrer aktiven Form aus 2 nicht-homologen α - und β -Untereinheiten. Diese bilden Heterodimere, Tetramere, und bei *Rhodococcus rhodochrous* J1 wurden sogar Decamere nachgewiesen. Die α - und β -Untereinheiten besitzen ungefähr die gleiche Größe, haben aber sonst keine Ähnlichkeiten untereinander.

Nitrilhydratasen sind Metalloproteine die Fe^{3+} oder Co^{3+} -Ionen enthalten (Bunch A. W. (1998), Nitriles, in: Biotechnology, Volume 8a, Biotransformations I, Chapter 6, Eds.: Rehm HJ, Reed G, Wiley-VCH, p. 277-324; Shearer J, Kung IY, Lovell S, Kaminsky W, Kovacs JA (2001) Why is there a "inert" metal center in the active site of nitrile hydratase? Reactivity and ligand dissociation from a five-coordinate $\text{Co}(\text{III})$ nitrile hydratase model. J Am Chem Soc

123: 463-468; Kobayashi M, Shimizu S (2000) Nitrile hydrolases. Current Opinion in Chemical Biology 4: 95-102).

Eine der größten Herausforderungen bisher ist die heterologe Darstellung von Nitrilhydratases in einem geeigneten Wirt, bevorzugt in *E. coli*. Dieses Gram negative Bakterium ist bekannt für seine hohen Expressionsraten heterologer Proteine. Ein weiterer Vorteil ist die Ausbeute an Biomasse in Hoch-Zelldichte-Fermentationen mit *E. coli*. Hierbei können Produktivitäten von über 100 g Biotrockenmasse (BTM) in 24 bis 44 Stunden erreicht werden (Lee SY (1996) High cell-density culture of *Escherichia coli*. TIBTECH 14:98-105; Riesenberger D, Guthke R (1999) High-cell-density cultivation of microorganisms. Appl Microbiol Biotechnol 51:422-430).

Die meisten Nitrilhydratase-Sequenzen der α - und β -Untereinheit sind aus der Gattung *Rhodococcus* bekannt. Aber gerade die Expression der Nitrilhydratases aus dieser Gattung in *E. coli* war bisher nur unter besonderen Schwierigkeiten möglich (Ikehata O, Nishiyama M, Horinouchi S, Beppu T (1989) Primary structure of nitrile hydratase deduced from the nucleotide sequence of a *Rhodococcus* species and its expression in *Escherichia coli*. Eur J Biochem 181: 563-570).

In der Literatur sind Expressionssysteme für Nitrilhydratases beschrieben deren spezifische Aktivitäten zwischen 4,2 und 12,2 U/mg Gesamtprotein für Co-abhängige Nitrilhydratases aus *R. rhodochrous* J1 (Kobayashi M, Nishiyama M, Nagasawa T, Horinouchi S, Beppu T, Yamada H (1991) Cloning, nucleotide sequence and expression in *Escherichia coli* of two cobalt-containing nitrole hydratase genes from *Rhodococcus rhodochrous*. Biochim Biophys Acta 1129: 23-33) und 452 U/mg Gesamtprotein für eine eisen-abhängige Nitrilhydratase aus *Rhodococcus spec. N-771* liegen (Njori M, Yohda M, Odaka M, Matsushita Y, Tsujimura M, Yoshida T, Dohmae N, Takio K, Endo I (1999) Functional

expression of Nitrile hydratases in *E. coli*: Requirement of a nitrile hydratase activator and a post-translational modification of a ligand cysteine. *J Biochem* 125: 696-704), was ungefähr ca. 248 U/mg BTM (Biotrockenmasse) entspricht (Kalkulation nach Goodsell DS (1991) *Inside a cell*. *TIBS* 16: 203-206). Interessanterweise konnte die letztgenannte Aktivität mit Nitrilhydratases aus *R. erythropolis*, welche nahe verwandt sind mit *Rhodococcus spec. N-711*, mit ähnlichen Vektorsystemen und Anordnungen der Strukturgene nicht nachvollzogen werden. Es bestand daher immer noch ein Bedarf an Verfahren und Systemen welche es gestatten, die ins Auge gefassten Enzyme in für technische Maßstäbe ausreichender Art und Weise zur Verfügung zu stellen.

Bisher beschriebene Methoden des Screenings nach Nitrilhydratases beschränkten sich auf die Isolierung von Mikroorganismen, die eine entsprechende Enzymaktivität aufweisen. Diese Mikroorganismen wurden entweder aus bestehenden Stammsammlungen entnommen, oder in so genannten Anreicherungsmedien selektiv angezogen (Colquhoun JA, Heald SC, Li L, Tamaoka J, Kato C, Horikoshi K and Bull AT (1998a) Taxonomy and biotransformation activities of some deep-sea actinomycetes. *Extremophiles* 2: 269-277; Colquhoun JA, Mexson J, Goodfellow M, Ward AC, Horikoshi K and Bull AT (1998b) Novel rhodococci and other mycolate actinomycetes from the deep sea. *Antonie van Leeuwenhoek* 74: 27-40).

Nachteilig an diesen Methoden ist, dass mit diesen Screening-Verfahren hauptsächlich Mikroorganismen der Gattungen *Rhodococcus*, *Bacillus*, *Pseudomonas*, o. ä. gefunden wurden und die Diversität der entsprechenden Nitrilhydratasesen bezüglich Aktivität oder Substratspezifität damit auch eingeschränkt ist (Bunch A. W. (1998), Nitriles, in: *Biotechnology*, Volume 8a, *Biotransformations I*, Chapter 6, Eds.: Rehm H.J., Reed G., Wiley-VCH, p. 277-324; Cowan D, Cramp R, Pereira R, Graham D, Almathawa Q (1998) *Biochemistry and biotechnology of*

mesophilic and thermophilic nitrile metabolising enzymes. Extremophiles 2: 207-216; Yamada H, Kobayashi M (1996) Nitrile hydratase and its application to industrial production of acrylamide. Biosci Biotechnol Biochem 60: 1391-141).

Derzeitige Schätzungen gehen davon aus, dass in der Regel nur 0,01 - 1% der Mikroorganismen eines Habitats kultivierbar sind und damit einem Screening nach der oben beschriebenen Methode zugeführt werden können (Amann RI, Ludwig W, Schleifer KH (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol. Rev. 59: 143-169; Pace NR (1997) A molecular view of microbial diversity and the biosphere. Science 276: 734-740; Cowan DA (2000) Microbial genomes--the untapped resource. Trends Biotechnol. 18: 14-16). Die Bedeutung der "direkten Klonierung" von genomischer DNA möglichst aller Organismen einer Bodenprobe (das "Metagenom") und deren Zuführung in ein genetisches Screening in Form von Metagenom-Genbanken gewinnt daher für die Identifizierung neuer technischer Enzyme immer mehr an Bedeutung. Das genetische Screening nach Enzym-kodierenden Genen kann hierbei Sequenz-homolog auf Basis konservierter Sequenzmotive oder bei Vorliegen geeigneter Enzymtests / Indikatormedien auch Aktivitäts-homolog erfolgen (Lorenz P, Köhler B, Wolf M, Eck J, Zinke H (2000) Expression Cloning of Metagenome DNA from Soil. Biotechnol. 2000, Book of Abstr. Vol 2: 306).

Mittels eines PCR-basierten Screenings unter Verwendung degenerierter Primer konnten bereits Nitrilhydratassen aus metagenomischer DNA amplifiziert werden, wobei aber die Sequenzen naturgemäß sehr hohe Ähnlichkeiten (90 - 99%) zu bekannten Nitrilhydratassen aufwiesen (Precigou S, Goulas P, Duran R, (2001) Rapid and specific identification of nitrile hydratase encoding genes in soil samples by polymerase chain reaction, FEMS Microbiol. Letters 204:

155-161). Eine fundierte Beurteilung der in dieser zitierstelle verwendeten Primer ist nicht möglich, da die Autoren die Sequenz der degenerierten Primer und deren Degenerationgrad nicht offen legen. Die hohe Ähnlichkeit 5 der Sequenzen zu der Nitrilhydratasen aus *Rhodococcus rhodochrous* J1 H lässt vermuten, dass auch die Substratspezifität nicht deutlich anders ist als bei dem genannten Enzym.

Es bestand daher immer noch ein Bedarf an Verfahren und 10 Systemen welche es gestatten, weitere der ins Auge gefassten Enzyme in für technische Maßstäbe ausreichender Art und Weise zur Verfügung zu stellen.

Aufgabe der vorliegenden Erfindung war deshalb die Angabe eines weiteren Verfahrens zur Herstellung von 15 Nitrilhydratasen. Insbesondere sollte das Verfahren im Stande sein, Nitrilhydratasen zu identifizieren, die in so genannten nicht kultivierbaren Organismen vorkommen. Eine weitere Aufgabe der vorliegenden Erfindung bestand in der Herstellung von gegenüber dem Stand der Technik 20 verbesserten Nitrilhydratasen.

Diese und weitere nicht näher spezifizierte, sich jedoch aus dem Stand der Technik in naheliegender Weise ergebende Aufgaben werden durch die Angabe spezieller Primerbestandteile gemäß Anspruch 1 und deren Anwendung in einem Verfahren mit den Merkmalen des Anspruchs 2 gelöst. 25 Anspruch 3 bis 5 beziehen sich auf bevorzugte Ausführungsformen des gegenständlichen Verfahrens. Anspruch 6 schützt die erfindungsgemäßen Proteinsequenzen, Anspruch 7 ist auf die sie codierenden Nukleinsäuresequenzen gerichtet und Anspruch 8 betrifft die mit diesen 30 Nukleinsäuresequenzen ausgestatteten Expressionssysteme. Anspruch 9 befasst sich mit erfindungsgemäß hergestellten neuen Nitrilhydratasen. Ansprüche 10 und 11 sind auf spezielle Verwendung gerichtet.

voraussetzung für das Auffinden neuer Nitrilhydrataser ist die Angabe von Nukleinsäuresequenzen, die im Stande sind, als Sonde für Nitrilhydratase-Gene in DNA-Metagenombibleotheken zu dienen. Durch die Angabe von 5 degenerierten Primerbestandteilen aus der Gruppe bestehend aus

A-01f : gcsmrsgcstgg (Seq. ID NO. 1)
B-01f : ggsctscsccc (Seq. ID NO. 2)
B-01r : ggsggsagscc (Seq. ID NO. 3)
10 C-01r : ggncgcwbsgg (Seq. ID NO. 4)
A-01f : gcnmrrgcntgg (Seq. ID NO. 5)
B-01f : ggnytnccncc (Seq. ID NO. 6)
B-01r : ggnggnarncc (Seq. ID NO. 7)
15 C-01r : gwngwrtccca (Seq. ID NO. 8)
A-01f : gcntggrynga (Seq. ID NO. 9)
B-01f : ggnytscncc (Seq. ID NO. 10)
B-01r : ggnggsarncc (Seq. ID NO. 11)
C-01r : swnswrtccca (Seq. ID NO. 12)

20 gelangt der Fachmann völlig überraschend dafür aber nicht minder vorteilhaft zu besonderen Nukleinsäuresequenzen, die helfen, spezielle Sonden für das Fischen nach Nitrilhydratase-Genen in DNA-Metagenombibliotheken zu erstellen. Es sind dies degenerierte Nukleinsäuresequenzen, 25 die auf der einen Seite spezifisch genug sind, nur Nitrilhydratase-Gene ausfindig zu machen, auf der andern Seite jedoch so unspezifisch sind, dass möglichst alle vorhandenen Nitrilhydratase-Gene erfasst werden. Deren Herstellung war zur Zeit der Erfindung aus dem Stand der 30 Technik in naheliegenderweise nicht herleitbar.

Gegenstand der vorliegenden Erfindung war demnach auch ein Verfahren zur Herstellung von Proteinsequenzen, welche notwendig zum Aufbau der Aktivität einer Nitrilhydratase sind, dergestalt, dass man
35 a) eine DNA-Metagenombibleothek eines Habitats erstellt,
b) diese mit mindestens jeweils einem forward(f)- und einem

reverse(r)-Primer aufweisend eine degenerierte Nukleinsäuresequenz gemäß den Sequenzen 1 bis 12 kontaktiert,

c) eine PCR hiermit durchführt,

5 d) aus den erhaltenen Teilsequenzen die Vollängensequenzen der Nukleinsäuren codierend für Proteinsequenzen notwendig zum Aufbau der Aktivität einer Nitrilhydratase generiert und

e) diese in einen Gastorganismus kloniert und exprimiert.

10 Wie eingangs schon angemerkt, bestehen Nitrilhydratassen zumindest aus zwei verschiedenen Untereinheiten (α - und β -Untereinheit). Es kann jedoch sein, dass neben diesen zwei Untereinheiten weitere Proteinsequenzen notwendig sind, um eine Aktivität der Nitrilhydratassen zu gewährleisten. Unter 15 Umständen ist das Vorhandensein bestimmter so genannter putativer „Aktivatoren“ (z. B. Faltungsproteine etc.) notwendig, damit entsprechende Aktivitäten von Nitrilhydratassen zu Stande kommen. Die für diese Aktivatoren codierenden Nukleinsäuresequenzen befinden sich 20 häufig in unmittelbarer Nähe der Nukleinsäuresequenzen kodierend für die entsprechenden Nitrilhydratase-Untereinheiten. Es ist demnach möglich, durch das Fischen nach Nukleinsäuresequenzen kodierend für Nitrilhydratassen alle die für die Aktivität der Nitrilhydratassen benötigten 25 Proteinsequenzen mitzuerfassen. Laut erfindungsgemäßem Verfahren geht man von der Erstellung einer DNA-Metagenombibliothek eines bestimmten Habitats aus. Wie diese Bibliothek erstellt wird, ist dem Fachmann geläufig (Knietsch, AW, Tanja; BS; Henne, ADR (2003) Metagenomes of 30 Complex Microbial Consortia Derived from Different Soils as Sources for Novel Genes Conferring Formation of Carbonyls from Short-Chain Polyols on *Escherichia coli*. Journal of Molecular Microbiology and Biotechnology 5(1): 46-56; Rondon, MR; August, PR; Bettermann AD; Brady, SF; Grossman, T H; Liles, MR; Loiacono, KA; Lynch, BA; MacNeil, IA; 35 Mino, r C; Tiong, CL; Gilman, M; Osburne, MS; Clardy, J;

Handelsman, J; Goodman, RM (2000) Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Applied and environmental microbiology* 66(6):2541-7). Diese wird im folgenden mit Primern versetzt, welche die erfindungsgemäß degenerierten Nukleinsäuresequenzen (Seq. ID No. 1-12) aufweisen. Anschließend wird eine PC-Reaktion durchgeführt, wobei Teilsequenzen der Nukleinsäuresequenzen kodierend für Nitrohydratase-Untereinheiten erhalten werden. Ausgehend von diesen Teilsequenzen kann der Fachmann mittels Methoden des Standes der Technik auf die entsprechenden Vollängensequenzen der entsprechenden Nukleinsäuresequenzen schließen (Schloss, PD; Handelsman, J (2003) Biotechnological prospects from metagenomics. 15 *Current Opinion in Biotechnology*, 14(3): 303-310; Rondon, MR; August, PR; Bettermann AD; Brady, SF; Grossman, T H; Liles, MR; Loiacono, KA; Lynch, BA; MacNeil, IA; Minor C; Tiong, CL; Gilman, M; Osburne, MS; Clardy, J; Handelsman, J; Goodman, RM (2000) Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Applied and environmental microbiology* 66(6):2541-7). Abschließend werden die gefundenen Nukleinsäuresequenzen in bestimmten Expressionssystemen rekombinant hergestellt. Dies ist dem Fachmann ebenfalls geläufig (Lit. s.u.).

20 In einer bevorzugten Ausführungsform werden die erfindungsgemäß degenerierten Nukleinsäuresequenzen (Seq. ID No. 1-12) im vorliegenden Verfahren der Gestalt eingesetzt, dass man jeweils Primerpaare aus Primern aufweisend die Nukleinsäuresequenzen A-01f (Seq. ID No. 1, 5, 9) und B-01r (Seq. ID No. 3, 7, 11) oder C-01r (Seq. ID No. 4, 8, 12) sowie B-01f (Seq. ID No. 2, 6, 10) und C-01r (Seq. ID No. 4, 8, 12) bei der PCR einsetzt. Unter diesen 25 Kombinationen werden die Nukleinsäuresequenzen kodierend für die Proteinsequenzen, welche notwendig für die Aktivität von Nitrohydratase sind, in bevorzugter und

effizienter Art und Weise detektiert.

Weiterhin ist bevorzugt, wenn den oben dargestellten degenerierten Primerbestandteilen gewisse andere Nukleinsäuresequenzen (z.B. „Stabilisierungsbereiche“)

5 vorgeschaltet sind (Kwok S, Chang SY, Sninsky JJ, Wang A, 1995, "Design and use of mismatched and degenerate primers" In: "PCR Primer, A laboratory Manual" Dieffenbach CW & Dveksler GS (Editors), Cold Spring Habor Laboratory Press, pp143-155; Compton T, 1990, "Degenerate Primers for DNA 10 Amplification" In: "PCR Protocols, A Guide to Methods and Applications", Innis MA, Gelfand DH, Sninsky JJ, White TJ (Editors) Academic Press, San Diego, pp39-45). In diesem Fall bestehen die bei der PC-Reaktion eingesetzten Primer aus degenerierten Nukleinsäuresequenzen der oben 15 dargestellten Art (Sequenzen ID No. 1 bis 12) und den in den Sequenzen mit den ID No. 13 bis 23 erwähnten Nukleinsäuresequenzen. Ganz besonders bevorzugt ist daher ein Verfahren, bei dem den degenerierten Nukleinsäuresequenzen (Seq. ID No. 1-12) solche ausgewählt 20 aus der Gruppe bestehend aus:

	GCCAAGGTCGTC	(Seq. ID NO. 13)
	GGCCGGTCCTG	(Seq. ID NO. 14)
	TCCTTGTACCAGGTC	(Seq. ID NO. 15)
	GCCCGGCC	(Seq. ID NO. 16)
	GGCGCTAATGTTGTT	(Seq. ID NO. 17)
25	TGGCCGGTTCTG	(Seq. ID NO. 18)
	CAAATTCTTATACCAAGTC	(Seq. ID NO. 19)
	CCATATATCGCATTTCAGCT	(Seq. ID NO. 20)
	GGTCGTGGCCAAG	(Seq. ID NO. 21)
30	GGCCGGTCCTG	(Seq. ID NO. 22)
	TCCTTGTACCAGGTC	(Seq. ID NO. 23)
	GCGCATTTCGGCG	(Seq. ID NO. 234)

vorangestellt sind. Diese Sequenzen sind ebenfalls aus konservierten Regionen von Nitrilhydratasen abgeleitet und 35 der Kodon-Verwendung von Organismen mit unterschiedlichem GC-Gehalt angepasst.

Ganz besonders vorteilhaft ist daher ein wie eingangs dargestelltes Verfahren, bei dem Primer eingesetzt werden, die ausgewählt sind aus der Gruppe bestehend aus

	GCCAAGGTCGTC <u>gcsmrsgcstgg</u>	(Seq. ID NO. 25)
5	GGCCGGTCCT <u>Gggsctscscc</u>	(Seq. ID NO. 26)
	TCCTTGTACCAGGT <u>Cggsggsagscc</u>	(Seq. ID NO. 27)
	GCCCGCC <u>ggncgcwbsgg</u>	(Seq. ID NO. 28)
	GGCGCTAAAGTT <u>Tgcnmrrgcntgg</u>	(Seq. ID NO. 29)
	TGGCCGGTCT <u>Gggnytnccncc</u>	(Seq. ID NO. 30)
10	CAAATTCTT <u>TATACCAAGTCggnggnarncc</u>	(Seq. ID NO. 31)
	CCATATATCG <u>CATTTCAGCTgwngwrtccca</u>	(Seq. ID NO. 32)
	GGT <u>CGTGGCCAAGgcntggrynga</u>	(Seq. ID NO. 33)
	GGCCGGTCCT <u>Gggnytsccncc</u>	(Seq. ID NO. 34)
	TCCTTGTACCAGGT <u>Cggnggsarncc</u>	(Seq. ID NO. 35)
15	GCGCATT <u>TCGGCGswnswrtccca</u>	(Seq. ID NO. 36).

Mit Hilfe dieser Primer konnten in DNA-Metagenombibleotheken Nukleinsäuresequenzen kodierend für Nitrilhydratasen sowie weitere Gene für putative „Aktivatoren“ nachgewiesen werden.

20 Demgemäß bildet einen nächsten Gegenstand der vorliegenden Erfindung die Proteinsequenzen, welche notwendig zum Aufbau der Aktivität einer Nitrilhydratase sind, wobei diese weniger als 100% Homologie, bevorzugt weniger als 97%, weiter bevorzugt weniger als 96%, noch weiter bevorzugt weniger als 95%, mehr bevorzugt weniger als 90%, ganz bevorzugt weniger als 85% und äußerst bevorzugt weniger als 80% Homologie, auf Aminosäureebene zu derartigen bekannten Proteinsequenzen besitzen und wobei die sie codierenden Nukleinsäuresequenzen aus Teilsequenzen generiert werden, die unter stringenten Bedingungen mit den erfindungsgemäßen Primern aufweisend die Nukleinsäuresequenzen mit den Sequenz ID No. 1 bis 12 ein positives Hybridisierungssignal geben.

25 Die positive Hybridisierung ist Voraussetzung dafür, dass entsprechende Nukleinsäuresequenzen mittels des PC-Reaktion

30

35

basierten Screenings gefunden werden können. Aus diesem Nukleinsäuresequenzen können dann nach Verfahren, die dem Fachmann geläufig sind, die entsprechenden rekombinanten Proteinsequenzen erhalten werden.

5 Durch solche rekombinante Techniken gelangt man zu Organismen, welche in der Lage sind, die betrachtete Proteinsequenz in für einen technischen Prozeß ausreichender Menge zur Verfügung zu stellen. Die Herstellung der erfundungsgemäßen rec-Proteinsequenzen 10 erfolgt nach dem Fachmann bekannten gentechnologischen verfahren (Sambrook, J.; Fritsch, E. F. und Maniatis, T. (1989), Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York; Balbas, P. und Bolivar, F. (1990), Design and construction of 15 expression plasmid vectors in E.coli, Methods Enzymol. 185, 14-37; Rodriguez, R.L. und Denhardt, D. T (eds) (1988), Vectors: a survey of molecular cloning vectors and their uses, 205-225, Butterworth, Stoneham). Bezuglich der allgemeinen Vorgehensweise (PCR, Klonierung, Expression 20 etc.) sei auch auf folgende Literatur und das dort zitierte verwiesen: Universal GenomeWalker™ Kit User Manual, Clontech, 3/2000 und dort zitierte Literatur; Triglia T.; Peterson, M. G. und Kemp, D.J. (1988), A procedure for in vitro amplification of DNA segments that lie outside the 25 boundaries of known sequences, Nucleic Acids Res. 16, 8186; Sambrook, J.; Fritsch, E. F. und Maniatis, T. (1989), Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York; Rodriguez, R.L. und Denhardt, D. T (eds) (1988), Vectors: a survey of 30 molecular cloning vectors and their uses, Butterworth, Stoneham.

Einen nächsten Gegenstand der vorliegenden Erfindung bilden auch die entsprechenden Nukleinsäuresequenzen, welche für die eben gezeigten Proteinsequenzen kodieren. Es sind dies 35 mithin alle, die im Umfang der Degeneration des genetischen Codes für die gleiche Proteinsequenz kodieren. Ebenfalls

umfasst sind damit auch solche, welche eine Homologie auf Nukleinsäureebene zu den erfindungsgemäß gefundenen Nukleinsäuresequenzen von mindestens 70 Prozent besitzen, oder entsprechende Fragmente dieser Nukleinsäuresequenzen, 5 die wiederum für Proteinsequenzen codieren, die am Aufbau der Aktivität einer Nitrilhydratase beteiligt sind. Bevorzugt kodieren diese für Proteinsequenzen, die gegenüber den erfindungsgemäß gefundenen Proteinsequenzen verbessert sind.

10 Erfindungsgemäße Nukleinsäuresequenzen sind z. B. die der ungeraden Seq. ID No. 37 bis 85. Es ist mit den gefundenen Nukleinsäuresequenzen - wie oben dargestellt - möglich, die erfindungsgemäßen Proteinsequenzen aus schnell wachsenden Wirtsorganismen, z.B. *E. coli*, in hohen Ausbeuten zu 15 gewinnen.

Dies geschieht durch Einbau (Klonierung) der erfindungsgemäßen Nukleinsäuresequenzen in spezielle Expressionssysteme, mit deren Hilfe die entsprechenden Proteinsequenzen aus bevorzugten Wirtsorganismen in 20 rekombinanter Art und Weise gewonnen werden können. Einen nächsten Aspekt der vorliegenden Erfindung bildet deshalb ein (künstlich hergestelltes) Expressionssystem aufweisend eine oder mehrere der erfindungsgemäßen Nukleinsäuresequenzen. Als Expressionssystem kommen 25 prinzipiell alle dem Fachmann für diesen Zweck geläufigen Systeme in Frage. Vorzugsweise sind dies Plasmide oder Vektoren und Mikroorganismen.

Als Plasmide oder Vektoren kommen im Prinzip alle dem Fachmann für diesen Zweck zur Verfügung stehenden Ausführungsformen in Frage. Derartige Plasmide und Vektoren 30 können z. B. von Studier und Mitarbeiter (Studier, W. F.; Rosenberg A. H.; Dunn J. J.; Dubendorff J. W.; (1990), Use of the T7 RNA polymerase to direct expression of cloned genes, Methods Enzymol. 185, 61-89) oder den Broschüren der 35 Firmen Novagen, Promega, New England Biolabs, Clontech oder

Gibco BRL entnommen werden. Weiter bevorzugte Plasmide und Vektoren können gefunden werden in: Glover, D. M. (1985), DNA cloning: a practical approach, Vol. I-III, IRL Press Ltd., Oxford; Rodriguez, R.L. und Denhardt, D. T (eds)

5 (1988), Vectors: a survey of molecular cloning vectors and their uses, 179-204, Butterworth, Stoneham; Goeddel, D. V. (1990), Systems for heterologous gene expression, Methods Enzymol. 185, 3-7; Sambrook, J.; Fritsch, E. F. und Maniatis, T. (1989), Molecular cloning: a laboratory 10 manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York.

Plasmide, mit denen das die erfindungsgemäße Nukleinsäuresequenzen aufweisende Konstrukt in ganz bevorzugter Weise in den Wirtsorganismus kloniert werden kann, sind: pUC18 (Roche Biochemicals), pKK-177-3H (Roche Biochemicals), pBTac2 (Roche Biochemicals), pKK223-3 (Amersham Pharmacia Biotech), pKK-233-3 (Stratagene) oder pET (Novagen). Äußerst bevorzugt sind Plasmide der pET-Reihe.

20 Der rekombinante Mikroorganismus, in den die die erfindungsgemäßen Nukleinsäuresequenzen enthaltenen Plasmide bzw. Vektoren kloniert werden, dient wie gesagt zur Vermehrung und Gewinnung einer ausreichenden Menge der rekombinanten Proteinsequenz. Die Verfahren hierfür sind dem Fachmann wohlbekannt (Sambrook, J.; Fritsch, E. F. und Maniatis, T. (1989), Molecular cloning: a laboratory 25 manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York). Als Mikroorganismen können im Prinzip alle dem Fachmann für diesen Zweck in Frage kommenden Organismen wie z.B. Hefen wie *Hansenula polymorpha*, *Pichia* sp., *Saccharomyces cerevisiae*, Prokaryonten, wie *E. coli*, *Bacillus subtilis* oder Eukaryonten, wie Säugerzellen, Insektenzellen oder Pflanzenzellen herangezogen werden. Vorzugsweise sind *E. coli*-Stämme für diesen Zweck zu 30 benutzen. Ganz besonders bevorzugt sind: *E. coli* XL1 Blue, NM 522, JM101, JM109, JM105, RR1, DH5 α , TOP 10 $^+$ oder HB101, 35

BL21, BL21 codon plus RIL, BL21 (DE3), oder BL21 (DE3) codon plus RIL.

Weiterhin eignen sich die erfindungsgemäßen (geraden Seq. ID No. 37 bis 85) und darüber hinaus weiter verbesserten 5 Nukleinsäuresequenzen, die für den Aufbau der Aktivität einer Nitrilhydratase notwendigen Proteinsequenzen codieren, bevorzugt zur Herstellung von sogenannten 10 Ganzzellkatalysatoren. Derartige Ganzzellkatalysatoren sind prinzipiell rekombinante Mikroorganismen wie die eben angesprochenen. Diese weisen jedoch neben den klonierten 15 Genen kodierend für eine Nitrilhydratase auch weitere für den Abbau von Nitrilen zu Säuren notwendigen Enzyme auf. Wie eingangs erläutert sind dies Enzyme, die eine Amidase-Aktivität aufweisen.

Als rekombinante Mikroorganismen der oben angesprochenen 20 Art gelten deshalb auch Ganzzellkatalysatoren aufweisend mindestens ein kloniertes Gen für eine Proteinsequenz mit (D- oder L-) Amidaseaktivität und klonierte Gene kodierend für eine aktive Nitrilhydratase. Optional kann der 25 Ganzzellkatalysator weitere Nukleinsäuresequenzen aufweisen, die für Enzyme kodieren, welche für den Abbau einer Nitril-Funktion in eine Säure-Funktion vorteilhaft sind. Es sind dies vor allem Enzyme ausgewählt aus der Gruppe der Proteinsequenzen mit α -Aminonitril-Racemaseaktivität, mit Cyanhydrin-Racemaseaktivität, mit α -Hydroxycarbonsäure-Racemaseaktivität oder mit (α - bzw. β)-Aminosäureamid-Racemaseaktivität.

Der erfindungsgemäße Ganzzellkatalysator produziert neben 30 den erfindungsgemäßen Proteinsequenzen, die zum Aufbau einer Nitrilhydratase-Aktivität notwendig sind, vorzugsweise eine Proteinsequenz mit L-Amidaseaktivität aus Rhizobium, vorzugsweise *R. huautlense* DSM 14983 (WO2004/005517) oder mit D-Amidaseaktivität, z. B. die aus *Variovorax* (EP 1318193).

35 Entsprechende Racemaser sind z.B. aus *Pseudomonas putida* und *Rhodococcus* sp. bekannt (Godtfredsen, S. E.; Clausen,

K.; Ingvorsen, K.; Hermes, H. F.; Van Balken, J. A.; Meijer, E. M. (1989, EP 0 307 023; WO 8 901 525). Weitere Aminosäureamid-Racemase bei *Klebsiella oxytoca* sind beschrieben bei Hermes und Mitarbeiter (Hermes, H. F. M.; 5 Peeters, W. P.; Peters, P. J. (1990), EP 0 383 403), sowie in *Agrobacterium rhizogenes* und *Ochrobacterium anthropi* (Boesten, W. H. J.; Raemakers-Franken, P. C.; Sonke, T.; Euverink, G. J. W.; WO 03106691). Der Vorteil des Einsatzes von entsprechenden Racematen liegt in der Tatsache 10 begründet, dass ein racemisches Nitril zu 100% in die entsprechende enantiomerenangereicherte Säure umgewandelt werden kann.

Vorzugsweise wird als Ganzzellkatalysator ein Organismus wie in der DE10155928 genannt als Wirtsorganismus 15 eingesetzt. Der Vorteil eines derartigen Organismus ist die gleichzeitige Expression mehrerer Enzymsysteme, womit nur noch ein rec-Organismus für die Reaktion von einem leicht herstellbaren Nitril bzw. Cyanhydrin oder α -Aminonitril 20 zur entsprechenden enantiomerenangereicherten Säure angezogen werden muß.

Um die Expression der ins Auge gefassten Nukleinsäuresequenzen im Hinblick auf die Umsetzungsgeschwindigkeiten der durch sie codierten Proteinsequenzen (Enzyme) abzustimmen, können die 25 entsprechenden Nukleinsäuresequenzen auf unterschiedlichen Plasmiden mit unterschiedlichen Kopienzahlen untergebracht und/oder unterschiedlich starke Promotoren für eine unterschiedlich starke Expression der Nukleinsäuresequenzen verwendet werden. Bei derart abgestimmten Enzymsystemen 30 tritt vorteilhafterweise eine Akkumulation einer ggf. inhibierend wirkenden Zwischenverbindung nicht auf und die betrachtete Reaktion kann in einer optimalen Gesamtgeschwindigkeit ablaufen. Dies ist dem Fachmann jedoch hinlänglich bekannt (Gellissen, G.; Piontek, M.; Dahlems, U.; Jenzelewski, V.; Gavagan, J. W.; DiCosimo, R.; 35 Anton, D. L.; Janowicz, Z. A. (1996), Recombinant Hansenula polymorpha as a biocatalyst. Coexpression of the spinach

glycolate oxidase (GO) and the *S. cerevisiae* catalase T (CTT1) gene, *Appl. Microbiol. Biotechnol.* 46, 46-54; Farwick, M.; London, M.; Dohmen, J.; Dahlems, U.; Gellissen, G.; Strasser, A. W.; DE19920712).

5 Mit dem vorliegenden Verfahren ist es demnach möglich, die α- und β-Untereinheiten von Nitrilhydratasen ausgehend von DNA-Metagenombibliotheiken herzustellen. Demgemäß bildet einen weiteren Gegenstand der vorliegenden Erfindung die Nitrilhydratasen aufweisend die erfindungsgemäßen Proteinsequenzen für α-Untereinheiten und β-Untereinheiten von Nitrilhydratase, welche sich aus den durch dieses Verfahren zugänglichen Nukleinsäuresequenzen kodierend für die erfindungsgemäßen α- und β-Untereinheiten herstellen lassen. Wie in den Beispielen gezeigt, ergeben sich dabei 15 auch aktiven Nitrilhydratasen, wenn beliebige α-Untereinheiten mit beliebigen β-Untereinheiten kombiniert werden. Dadurch ist es möglich, die Diversifiziertheit von möglichen Nitrilhydratasen weiter zu steigern.

Weitere Gegenstände der vorliegenden Erfindung beziehen 20 sich auf die Verwendung der durch das erfindungsgemäße Verfahren hergestellten Nukleinsäuresequenzen zur Erzeugung von verbesserten Proteinsequenzen notwendig zum Aufbau der Aktivität einer Nitrilhydratase. Wie der Fachmann bei der Verbesserung von Proteinsequenzen auf Basis der Änderung 25 von Nukleinsäuresequenz vorgeht, ist allgemein bekannt. Dies geschieht im allgemeinen durch Anwendung von Mutagenese-Methoden. Als Mutagenese-Methoden kommen alle dem Fachmann für diesen Zweck zur Verfügung stehenden Methoden in Frage. Insbesondere sind dies die 30 Sättigungsmutagenese, die Random-Mutagenesis, in vitro-Rekombinations-Methoden sowie Site-Directed-Mutagenesis (Eigen, M. und Gardiner, W. (1984), Evolutionary molecular engineering based on RNA replication, *Pure Appl. Chem.* 56, 967-978; Chen, K. und Arnold, F. (1991), Enzyme engineering 35 for nonaqueous solvents: random mutagenesis to enhance

activity of subtilisin E in polar organic media. Bio/Technology 9, 1073-1077; Horwitz, M. und Loeb, L. (1986), Promoters Selected From Random DNA-Sequences, Proc Natl Acad Sci USA 83, 7405-7409; Dube, D. und L. Loeb (1989), Mutants Generated By The Insertion Of Random Oligonucleotides Into The Active-Site Of The Beta-Lactamase Gene, Biochemistry 28, 5703-5707; Stemmer, P.C. (1994), Rapid evolution of a protein *in vitro* by DNA shuffling, Nature 370, 389-391 und Stemmer, P.C. (1994), DNA shuffling by random fragmentation and reassembly: *In vitro* recombination for molecular evolution. Proc Natl Acad Sci USA 91, 10747-10751).

Die so erhaltenen neuen Nukleinsäuresequenzen werden nach den oben angegebenen Methoden in einen Wirtsorganismus 15 kloniert, exprimiert und die so hergestellten Proteinsequenzen mit geeigneten Screening-Methoden detektiert und anschließend isoliert. Zur Detektion sind grundlegend alle möglichen Nachweisreaktionen für die gebildeten Moleküle geeignet. Nitrilhydratase-aktivitäten können in einem gekoppelten enzymatischen Test mit Amidasen 20 nachgewiesen werden, wobei Ammonium als Nebenprodukt gebildet wird. Für deren Nachweis eignen sich grundlegend alle möglichen Nachweisreaktionen für Ammoniak bzw. Ammoniumionen wie Nessler-Reagenz (Vogel, A., I., (1989) Vogel's textbook of quantitative chemical analysis, John Wiley & Sons, Inc., 5th ed., 679-698, New York) die 25 Indophenolreaktion auch Berthelot'sche Reaktion genannt (Wagner, R., (1969) Neue Aspekte zur Stickstoffanalytik in der Wasserchemie, Vom Wasser, VCH-Verlag, Bd. 36, 263-318, Weinheim) insbesondere die enzymatische Bestimmung mittels der Glutamat-Dehydrogenase (Bergmeyer, H., U., und Beutler, H.-O. (1985) Ammonia, in: Methods of Enzymatic Analysis, VCH-Verlag, 3rd Edition, Vol. 8: 454-461, Weinheim) aber 30 auch der Nachweis mit Ammonium-sensitiven Elektroden. Weiterhin dienen HPLC-Methoden zum Nachweis von Aminosäuren 35 wie z.B. ein Derivativ-Verfahren auf der Basis von O-Pthaldialdehyd und N-Isobutyryl-Cystein zur

Enantiomerentrennung von Aminosäuren (Brückner, H., Wittner R., und Godel H., (1991) Fully automated high-performance liquid chromatographic separation of DL-amino acids derivatized with o-Phthaldialdehyde together with N-isopropyl-cysteine. Application to food samples, Anal. Biochem. 144, 204-206). Das direkt durch die Nitrilhydratasereaktion gebildete Amid kann ebenfalls über HPLC-Methoden (z. B. inverse_Phase) nachgewiesen werden.

In einer letzten Ausgestaltung der vorliegenden Erfindung 10 bezieht sich diese auf die Verwendung der erfindungsgemäßen Nitrilhydratasen zur Herstellung von organischen Säureamiden und Säuren, insbesondere enantiomerangereicherten α -Hydroxysäuren bzw. α -Aminosäuren.

15 Für die Anwendung kann die ins Auge gefasste Nitrilhydratase in freier Form als homogen aufgereinigte Verbindungen oder als rekombinant hergestelltes Enzym verwendet werden. Weiterhin kann das Enzym auch als Bestandteil eines intakten Gastorganismus eingesetzt werden 20 oder in Verbindung mit der aufgeschlossenen und beliebig hoch aufgereinigten Zellmasse des Wirtsorganismus.

Möglich ist ebenfalls die Verwendung des Enzyms in 25 immobilisierter Form (Sharma B. P.; Bailey L. F. und Messing R. A. (1982), Immobilisierte Biomaterialien - Techniken und Anwendungen, Angew. Chem. 94, 836-852).

Vorteilhafterweise erfolgt die Immobilisierung durch 30 Lyophilisation (Paradkar, V. M.; Dordick, J. S. (1994), Aqueous-Like Activity of α -Chymotrypsin Dissolved in Nearly Anhydrous Organic Solvents, J. Am. Chem. Soc. 116, 5009-5010; Mori, T.; Okahata, Y. (1997), A variety of lipi-coated glycoside hydrolases as effective glycosyl transfer catalysts in homogeneous organic solvents, Tetrahedron Lett. 38, 1971-1974; Otamiri, M.; Adlercreutz, P.; Matthiasson, B. (1992), Complex formation between 35 chymotrypsin and ethyl cellulose as a means to solubilize the enzyme in active form in toluene, Biocatalysis 6, 291-

305). Ganz besonders bevorzugt ist die Lyophilisation in Gegenwart von oberflächenaktiven Substanzen, wie Aerosol OT oder Polyvinylpyrrolidon oder Polyethylenglycol (PEG) oder Brij 52 (Diethylenglycol-mono-cetylether) (Kamiya, N.; 5 Okazaki, S.-Y.; Goto, M. (1997), Surfactant-horseradish peroxidase complex catalytically active in anhydrous benzene, *Biotechnol. Tech.* 11, 375-378). Äußerst bevorzugt ist die Immobilisierung an Eupergit®, insbesondere Eupergit C® und Eupergit 250L® (Röhm) 10 (Eupergit.RTM. C, a carrier for immobilization of enzymes of industrial potential. Katchalski-Katzir, E.; Kraemer, D. M. *Journal of Molecular Catalysis B: Enzymatic* (2000), 10(1-3), 157-176).

Gleichfalls bevorzugt ist die Immobilisierung an Ni-NTA in 15 Kombination mit dem His-Tag (Hexa-Histidin) ergänzten Polypeptid (Purification of proteins using polyhistidine affinity tags. Bornhorst, Joshua A.; Falke, Joseph J. *Methods in Enzymology* (2000), 326, 245-254).

Die Verwendung als CLECs ist ebenfalls denkbar (St. Clair, 20 N.; Wang, Y.-F.; Margolin, A. L. (2000), Cofactor-bound cross-linked enzyme crystals (CLEC) of alcohol dehydrogenase, *Angew. Chem. Int. Ed.* 39, 380-383).

Durch diese Maßnahmen kann es gelingen aus Polypeptiden, welche durch organische Solventien instabil werden, solche 25 zu generieren, die in Gemischen von wässrigen und organischen Lösungsmitteln bzw. ganz in Organik arbeiten können.

Die vorliegende Erfindung beschreibt die Gewinnung von Nitrilhydratasen und deren korrespondierende Gene durch die 30 Erschließung der nicht-kultivierten mikrobiellen Diversität verschiedener Habitate mittels molekulargenetischer Methoden. Durch Verwendung degenerierter Primer werden beim PCR-basierten Durchmustern von DNA-Metagenombibliotheken Nitrilhydratase-Gene identifiziert und die Teilsequenz der 35 so erhaltenen PCR-Produkte aufgeklärt. In einem Folgeschritt wird die vollständige DNA-Sequenz der Gene

bestimmt, um nach Klonierung und heterologer Expression Enzymproben für Aktivitätsprofilierungen und Anwendungsuntersuchungen bereitzustellen.

Durch die rationale Auswahl von Bodenproben, die Nitrile enthalten könnten, für die Erstellung von Metagenom-Banken (DNA-Metagenombibliotheken) und die Fokussierung auf anwendungsrelevante Substrate im Rahmen von Anreicherungskulturen, kann eine Anreicherung von Nitrilumsetzenden Mikroorganismen erzielt werden. In jedem Falle liefert das genetische Screening in Metagenom-Banken ein 10 Pool an entsprechenden Nitrilhydratase-Genen zur nachfolgenden Expression, der aber auch als Grundlage für eine Enzymoptimierung durch gerichtete Evolution dienen kann.

15 Zur Identifizierung neuer Nitrilhydratasen wurden Metagenombibliotheken aus fünf unterschiedlichen Habitaten und Standorten (Grasland, Wald, sandiges Ökosystemen, Biofilm) mit mehr als 83.000 Klonen durchmustert. Der Aufbau solcher Metagenombibliotheken ist dem Fachmann geläufig (Schloss, PD; Handelsman, J (2003) 20 Biotechnological prospects from metagenomics. Current Opinion in Biotechnology, 14(3): 303-310; Knietsch, AW, Tanja; BS; Henne, ADR (2003) Metagenomes of Complex Microbial Consortia Derived from Different Soils as Sources 25 for Novel Genes Conferring Formation of Carbonyls from Short-Chain Polyols on *Escherichia coli*. Journal of Molecular Microbiology and Biotechnology 5(1): 46-56; Rondon, MR; August, PR; Bettermann AD; Brady, SF; Grossman, T H; Liles, MR; Loiacono, KA; Lynch, BA; MacNeil, IA; 30 Mino, r C; Tiong, CL; Gilman, M; Osburne, MS; Clardy, J; Handelsman, J; Goodman, RM (2000) Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. Applied and environmental microbiology 666): 2541-7). Diese Klone 35 enthalten zusammen etwa 3000 MBp DNA. Die DNA-

Metagenombibliotheken wurden mit Hilfe von degenerierten Oligonukleotiden, abgeleitet von konservierten Primärstrukturmotiven bekannter Nitrilhydratases, mittels eines PCR-Screenings nach neuen Nitrilhydratases 5 durchsucht. Als degenerierte Primer wurden solche gemäß den Sequenz ID No. 25 bis 36 benutzt.

Dabei können insbesondere jeweils die Primer mit dem Primerbestandteil A-01f mit den mit dem Primerbestandteilen B-01r (Größe des zu erwartenden PCR-produkts ca. 210 Bp) 10 oder C-01r (Größe des zu erwartenden PCR-produkts ca. 350 Bp) und die primer mit dem Primerbestandteile B-01f mit dem mit dem Primerbestandteil C-01r (Größe des zu erwartenden PCR-produkts ca. 180 Bp) kombiniert werden. Insbesondere 15 der degenerierte „Kern“-Teil der Primer (unterstrichen in der Sequenz ID No. 25 bis 36) ist für das Identifizieren neuer Nitrilhydratases von Bedeutung, während der nicht degenerierte Abschnitt variiert werden kann.

Die entsprechende PC-Reaktion kann nach bekannten Methoden durchgeführt werden. Bei der Verwendung besonderer 20 Polymerasen ist die PCR nach Maßgabe der Herstellerangaben abzuändern.

In allen 5 Metagenombibliotheken konnten Klone identifiziert werden, die die Gene für Nitrilhydratases tragen. Dem Fachmann sind Methoden bekannt, um mit der 25 Teilsequenz der Nitrilhydratase-Gene den das Gen tragenden Klon zu identifizieren und die Vollängensequenz aufzuklären (Schloss, PD; Handelsman, J (2003) Biotechnological prospects from metagenomics. Current Opinion in Biotechnology, 14(3): 303-310; Duran, R; 30 Nishiyama, M; Horinouchi, S; Beppy, T (1993) Characterization of nitrile hydratase genes cloned by DNA screening from *Rhodococcus erythropolis*. Biosci Biotech Biochem 57(8): 1323-1328).

So konnte die Vollängensequenzen (Nukleinsäuresequenzen) für insgesamt 12 Gene von α -Untereinheiten sowie für 10 Gene von β -Untereinheiten bestimmt werden (Abbildung 1). Darüber hinaus wurde die Sequenz von drei Genen bestimmt, 5 die für putative „Aktivatoren“ von Nitrilhydratassen kodieren könnten.

zur Darstellung der Nitrilhydratassen wurden die Gene der α - und β -Untereinheiten jeweils separat unter die Kontrolle des Promotors des Gens 10 des Phagen T7 gestellt. Dazu 10 wurden die Vektoren pET22b bzw. pET26b (Novagen, CN Bioscience, Inc.) verwendet (Abb. 5 und 6). Die Anwendung eines Zwei-Vektor-Expressionssystems erlaubte die Möglichkeit der einfachen Kombination der Nitrilhydratase-Untereinheiten verschiedener Enzyme. Entsprechende 15 Konstrukte wurden im Stamm *E. coli* BL21 (DE3) CodonPlus RIL (Novagen, CN Bioscience, Inc.) exprimiert. Die Zellen wurden bei 26°C unter Verwendung von LB-Medium inkubiert und bei Erreichen einer Zelldichte von O.D.580 = 1.0 mit 0,5 mM IPTG induziert. Dem vereinzelt auftretenden Problem 20 der Proteinaggregation nach Überexpression bei 37°C konnte durch Reduktion der Temperatur auf 26°C begegnet werden. Die Koexpression verschiedener Chaperone (Trigger-Faktor, GroEL/GroES und Dnak/DnaJ/GrpE) zeigte bei 26°C keinen zusätzlichen Effekt, der über den durch die 25 Temperaturreduktion erzielten hinausging.

Die Ausbeuten hinsichtlich der Aktivität gegen Benzonitril variierten erheblich (Abb. 2), wobei keine eindeutige Korrelation zwischen der dargestellten Proteinmenge und der erzeugten Aktivität nachzuweisen war. So war die Ausbeute 30 für den Klon mit der Nitrilhydratase M49bD9 (Seq. ID No. 49/71) mit 30 U/g BTM relativ hoch, obwohl in der SDS-PAGE-Analyse keine Proteine für die Nitrilhydratase nachweisbar waren. Demgegenüber steht der Klon M12K24 (Seq. ID No. 39/63), für den eine Aktivität von etwa 2,5 U/g BTM 35 bestimmt wurden, obwohl beide Untereinheiten relativ stark

überexprimiert werden konnten und ein beträchtlicher Anteil in der löslichen Fraktion zu finden war. Die Höhe der Aktivität hängt also stark von dem jeweiligen Enzym ab.

Für verschiedene Nitrilhydratasen wurden in unmittelbarer Nähe zu den Genen für die Untereinheiten des Enzyms ein weiterer, kleiner offener Leserahmen beschrieben, der für eine Proteinsequenz kodiert, die an der Aktivierung der Nitrilhydratase beteiligt zu sein scheint. Diese Proteine wurden als P12K (Seq. ID No. 81/83) bzw. P14K (Seq. ID NO. 85) bezeichnet, da ihr Molekulargewicht etwa 12 bzw. 14 kDa beträgt. Während einige Nitrilhydratasen auch ohne diese P12K-Homologen aktiv dargestellt werden konnten, war die Anwesenheit dieser Proteine für die aktive Expression anderer Nitrilhydratasen essentiell. Um den Einfluß von P12K-Homologen, deren Gene auch in drei Metagenom-Klonen gefunden wurden, auf die Expression der entsprechenden Nitrilhydratasen aber auch auf Nitrilhydratasen von anderen Klonen zu untersuchen, wurden zwei dieser Gene in den Vektor pBBR1MCS5 (Kovach et al., 1995, Four new derivatives of the broad host range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes, Gene 166: 175-176) kloniert und hier unter die Kontrolle des lac-Promotors gestellt. Die entsprechenden Konstrukte wurden mit pBBR5-P12K-M49bD9 und pBBR5-P12K-M3aG10 bezeichnet (Abb. 6 und 7).

Die Expression der Nitrilhydratase-Gene M49bD9 (Seq. ID No. 49/71) und M3aG10 (Seq. ID No. 57/79) in Gegenwart der entsprechenden P12K-Homologen (Seq. ID No. 83 - M49bD9; Seq. ID No. 85 - M3aG10) erfolgte im Stamm *E. coli* BL21 CodonPlus RIL bei 26°C. In beiden Fällen konnten die Untereinheiten der Nitrilhydratasen deutlich überexprimiert werden.

Die Aktivität der Nitrilhydratasen des Klons M49bD9 (Seq. ID No. 49/71) wurde durch die Anwesenheit des P12K-Homologen (Seq. ID No. 83) etwa um den Faktor 27 auf ca.

830 U/g BTM gesteigert (Tabelle 1). Für das Enzym des Klons M3aG10 (Seq. ID No. 57/79) konnte unter diesen Bedingungen zum erstenmal Aktivität (ca. 23 U/g BTM) nachgewiesen werden. Diese Ergebnisse belegen, dass die Anwesenheit der 5 P12K-Homologen (Seq. ID No. 85) für die Erhöhung der Aktivitätsausbeute entscheidend kann.

Tabelle 1: Aktivität von Nitrilhydratases nach Koexpression mit P12K-Homologen (Seq. ID No. 83 und 85)

Klon	Ohne P12K	Mit P12K
M49bD9 (α, β)	30 U/g BTM	826 U/g BTM
M3aG10	0 U/g BTM	23 U/g BTM

10 Die Kombination von Untereinheiten aus verschiedenen Metagenom-Klonen eröffnet die Möglichkeit Nitrilhydratases mit potentiell neuen Substratspezifitäten zu erzeugen. Durch die Kombination von α -Untereinheiten mit 15 verschiedenen β -Untereinheiten würde eine große Diversität an Kombinationen neuer Nitrilhydratases möglich werden. Bis dato ist eine solche Kombination von Untereinheiten aus nicht zusammengehörigen Nitrilhydratases in der Literatur nicht bekannt.

20 Um diese Möglichkeit zu evaluieren, wurden die α -Untereinheiten der Klone M73dC9 und M15aA6 (Seq. ID No. 59 und 45), für die keine β -Untereinheiten gefunden werden konnten, mit der β -Untereinheit des Klons M12K24 (Seq. ID No. 63) im Stamm *E. coli* BL21 codon plus RIL exprimiert.

25 Während sich bei der Expression des Paars von a-M73dC9a/b-M12K24 (Seq. ID No. 59 und 63) beide Untereinheiten etwa gleich stark darstellen ließen, scheint die α -Untereinheit des Klons M15aA6 (Seq. ID No. 45) besser exprimiert worden

zu sein als die β -Untereinheit des Klons M12K24 (Seq. ID No. 63).

Die Kombination der α -Untereinheit des Klons M73dC9 (Seq. ID No. 59) und der β -Untereinheit des Klons M12K24 (Seq. ID No. 63) führte überraschenderweise zur Bildung einer aktiven Nitrohydratase mit einer Aktivität von ca. 0,07 U/g BTM (Abb. 3). Dieses Ergebnis belegt, daß es prinzipiell möglich ist, durch Kombination von Nitrohydratase-Untereinheiten verschiedener Klone aktive Enzyme darzustellen. Dieses war zum Zeitpunkt der Erfindung aus dem Stand der Technik als solches nicht herleitbar.

Stringente Bedingungen: Der Ausdruck "unter stringenten Bedingungen" wird hierin wie bei Sambrook et al. (Sambrook, J.; Fritsch, E. F. und Maniatis, T. (1989), Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York) beschrieben, verstanden. Bevorzugt liegt eine stringente Hybridisierung gemäß der vorliegenden Erfindung vor, wenn nach Waschen für eine Stunde mit 1 x SSC (150 mM Natriumchlorid, 15 mM Natriumcitrat, pH 7.0) und 0,1 % SDS (Natriumdodecylsulfat) bei 50 °C, bevorzugt bei 55 °C, mehr bevorzugt bei 62 °C und am meisten bevorzugt bei 68 °C und mehr bevorzugt für 1 Stunde mit 0,2 x SSC und 0,1 % SDS bei 50 °C, bevorzugter bei 55 °C, mehr bevorzugt bei 62 °C und am meisten bevorzugt bei 68 °C noch ein positives Hybridisierungssignal beobachtet wird.

Unter optisch angereicherten (enantiomerenangereicherten, enantiomer angereicherten) Verbindungen wird im Rahmen der Erfindung das Vorliegen einer optischen Antipode im Gemisch mit der anderen in >50 mol-% verstanden.

Unter dem Begriff Nukleinsäuresequenzen werden alle Arten von einzelsträngiger und dazu komplementärer oder doppelsträngiger DNA (z.B. genomische, cDNA) als auch RNA (z.B. mRNA) oder Gemische derselben subsumiert.

verbesserte Nukleinsäuresequenzen kodieren für verbesserte Proteinsequenzen. Verbesserte Proteinsequenzen sind solche, bei denen eine Verbesserung im Hinblick auf die Aktivität und/oder Selektivität und/oder Stabilität gegenüber den Ursprungssequenzen zu verzeichnen ist. Erfindungsgemäß bedeutet dies, dass die Proteine aktiver und/oder selektiver bzw. weniger selektiv oder unter den verwendeten Reaktionsbedingungen stabiler sind. Während die Aktivität und die Stabilität der Proteine für die technische Anwendung naturgemäß möglichst hoch sein sollte, ist in Bezug auf die Selektivität dann von einer Verbesserung die Rede, wenn entweder die Substratselektivität abnimmt, die Enantioselektivität der Proteine jedoch gesteigert ist. Dies gilt ebenfalls für Proteine als Bestandteile von Nitrilhydratasen, sofern sie dem Enzym die entsprechenden verbesserten Eigenschaften verleihen helfen.

Von den beanspruchten Proteinsequenzen und den Nukleinsäuresequenzen werden erfindungsgemäß auch solche Sequenzen umfaßt, die eine Homologie (exclusive der natürlichen Degeneration) größer als 70% (in Bezug auf die Nukleinsäuresequenz) bzw. 80% (in Bezug auf die Proteinsequenzen), bevorzugt größer als 90%, 91%, 92%, 93% oder 94%, mehr bevorzugt größer als 95% oder 96% und besonders bevorzugt größer als 97%, 98% oder 99% zu einer dieser Sequenzen aufweisen, sofern die Wirkungsweise bzw. Zweck einer solchen Sequenz erhalten bleibt. Der Ausdruck "Homologie" (oder Identität) wie hierin verwendet, kann durch die Gleichung $H (\%) = [1 - V/X] \times 100$ definiert werden, worin H Homologie bedeutet, X die Gesamtzahl an Nukleobasen/Aminosäuren der Vergleichssequenz ist und V die Anzahl an unterschiedlichen Nukleobasen/Aminosäuren der zu betrachtenden Sequenz bezogen auf die Vergleichssequenz ist. Auf jeden Fall sind mit dem Begriff Nukleinsäuresequenzen, welche für Polypeptide codieren, alle Sequenzen umfaßt, die nach Maßgabe der Degeneration des genetischen Codes möglich erscheinen.

Beschreibung der Abbildungen:

Abbildung 1: Homologie unter den mittels eines genetischen Screenings gefundenen α - und β -Untereinheiten von Nitrilhydratasen.

5

Abbildung 2: Aktivität verschiedener Metagenom-Nitrilhydratasen gegen Benzonitil nach Expression in *E. coli* BL21 (DE3) codon plus RIL mit und ohne Koexpression des Triggerfaktors (n.b.: nicht bestimmt).

10

Abbildung 3: Aktivität von Nitrilhydratasen bei Kombination von Untereinheiten aus verschiedenen Metagenom-Klonen.

15 Abbildung 4: Die Vektorkarte zeigen die generelle Anordnung der α - Untereinheiten im Plasmid pET22 am Beispiel des Klones M49bD9.

20

Abbildung 5: Die Vektorkarte zeigen die generelle Anordnung der β - Untereinheiten im Plasmid pET26 am Beispiel des Klones M49bD9.

Abbildung 6: Die Vektorkarte zeigen die Anordnung des Proteins P12K des Klones M49bD9 im Plasmid pBBR5.

25

Abbildung 7: Die Vektorkarte zeigen die Anordnung des Proteins P12K des Klones M3aG10 im Plasmid pBBR5.

Experimenteller Teil:

Allgemeines PCR-Protokoll:

5 Kultivierung von Mikroorganismen

Die Kultivierung der E. coli Zellen und deren Aufbewahrung erfolgte nach Sambrook et al. (Sambrook, J.; Fritsch, E. F. und Maniatis, T. (1989), Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York).

PCR-Protokoll:

Ein typisches PCR-Protokoll ist im folgenden dargestellt, 15 wobei ein Anpassen des Protokolls bei Verwendung einer anderer Polymerase nach Maßgabe der Herstellerangaben erforderlich wird.

25 µl HotStarTaq-Mastermix, Qiagen
(2,5 U Polymerase, 200 µM dNTPs, 1x PCR-Puffer)

20 25 pmol Primer 01f

25 pmol Primer 01r

2 µl template-DNA (20 - 200 ng)

ad 50 µl A. dest

Programm:

Initiale Denaturierung: 1x 15 min 95°C

zyklenprogramm: 35x
$$\begin{cases} 1 \text{ min } 95^\circ\text{C} \\ 1 \text{ min } 55^\circ\text{C} - 69^\circ\text{C} \\ 1 \text{ min } 72^\circ\text{C} \end{cases}$$

5

Finale Elongation: 1x 7 min 72°C

Last Extension: 7min 72°C

Verdau mit Restriktionsenzymen

10 Die zu schneidende DNA wird mit 5 U Restriktionsenzym und den zugehörigen Puffer versehen und wenn nicht anders erforderlich bei 37°C inkubiert. Der Verdau chromosomaler DNA erfolgt mit 10 U Enzym. Die Inkubationsdauer beträgt 1,5 - 2,5 Stunden.

15

Behandlung mit alkalischer Phosphatase

Um zu verhindern, dass Vektoren, welche nur mit einer Restriktionsendonuklease geschnitten wurden, mit sich selbst religieren, wird mit Hilfe der alkalischen Phosphatase der am 5'-Ende überhängende Phosphatrest entfernt. Nur durch Insertion eines DNA-Fragments kann wieder zirkuläre DNA entstehen.

Der mit einer Restriktionsendonuklease geschnittene Vektor wird 15 min bei 65°C inkubiert, um die Restriktionsendonuklease abzustoppen. Anschließend wird der Dephosphorylierungspuffer zugegeben und mit 1 U alkalischer Phosphatase aus Schrimps wird 10 min bei 37°C inkubiert.

Das Enzym wird durch eine anschließende Gelelektrophorese von der Vektor-DNA abgetrennt.

Behandlung mit T4-DNA-Ligase

5 Für die Ligation werden Vektor und Insert im Verhältnis 1:3 eingesetzt. Das Volumen wird möglichst gering gewählt (7-20 µl) Der Ansatz wird in Ligationspuffer und in Anwesenheit von 1 U Ligase bei 16 °C über Nacht inkubiert.

10 Transformation

Zum Ligationsansatz werden 100 µl kompetente Zellen pipettiert und der Ansatz durch wiederholtes Aufziehen der Pipette gemischt. Nach 30 min Inkubation auf Eis wird ein Hitzeschockschritt bei 42 °C für 45 sec durchgeführt und 15 wieder 2 min auf Eis inkubiert. Es werden 120-900 µl SOC-Medium zugegeben und der Ansatz wird 45 min bei 37°C unter Agitation inkubiert. Anschließend wird der Ansatz ausplattiert und bei 37°C über Nacht inkubiert.

20 Expression von Metagenom-Nitrilhydratasen

Die Expression der Konstrukte mit T7-Promotoren erfolgte nach folgendem Protokoll:
25 50 ml LB_{amp100}-Medium mit 2 mM Fe-Citrat und jeweils 50 µg/ml Kanamycin & Ampicillin wurde 1%ig mit einer Übernachtkultur angeimpft. Nach Erreichen einer OD₆₀₀ von ca. 0,5 wurde mit 1 mM IPTG (Isopropylthiogalactosid) die Expression der Nitrilhydratasen induziert. Die Ernte der Zellen erfolgte ca. 24 Stunden nach Induktion bei 26°C.

Aktivitätsnachweis mit Benzonitril als Substrat

Die Biotransformation wurde im 10 ml Maßstab durchgeführt mit ca. 100 mg Biofeuchtmasse ($OD_{600} = 5$) im Kaliumphosphatpuffer (100 mM) pH 7,0. Die Inkubation 5 erfolgte bei 30°C und die Substratkonzentration betrug ca. 5 mM Benzonitril. Die Probennahme erfolgte alle 5 - 10 min über einen Zeitraum von maximal 1 Stunde. Das Probenvolumen betrug 100 μ l und die Reaktion wurde durch Zugabe von 10 μ l 50%ige Phosphorsäure gestoppt.

10 Die Konzentrationen von Benzonitril und Benzamid wurden dann mittels HPLC bestimmt:

Säule: RP18-Säule Phenomenex Hypersil ODS 5 μ (mit Vorsäule)

Fließmittel: 10 mM K₂HPO₄, (pH 2.3)

Flußrate: 1 ml / min

15 Wellenlänge: 202 nm

Injektionsvolumen: 20 μ l

Dauer HPLC Lauf: 12-15 min

Die Berechnung der Aktivität erfolgte über die Kalkulation von einem μ mol Umsatz nach einer Minute, was einem U (Unit entspricht). Spezifische Aktivitäten werden in U pro g BTM 20 oder mg Protein angegeben.

Patentansprüche:

1. Degenerierte Primerbestandteile aus der Gruppe bestehend aus

5 A-01f : gcsmrsgcstgg (Seq. ID NO. 1)
 B-01f : ggsctscsccc (Seq. ID NO. 2)
 B-01r : ggsggsagscc (Seq. ID NO. 3)
 C-01r : ggnncgcwbsgg (Seq. ID NO. 4)
 A-01f : gcnmrrgcntgg (Seq. ID NO. 5)
 B-01f : ggnytnccncc (Seq. ID NO. 6)
 B-01r : ggnggnarncc (Seq. ID NO. 7)
 C-01r : gwngwrtccca (Seq. ID NO. 8)
 A-01f : gcntggrynga (Seq. ID NO. 9)
 B-01f : ggnytsccncc (Seq. ID NO. 10)
 B-01r : ggnggsarncc (Seq. ID NO. 11)
 C-01r : swnswrtccca (Seq. ID NO. 12).

10 2. Verfahren zur Herstellung von Proteinsequenzen notwendig zum Aufbau der Aktivität einer Nitrilhydratase dergestalt, dass man

15 a) eine DNA-Metagenombibliothek eines Habitats erstellt,

20 b) diese mit mindestens jeweils einem forward(f)- und einem reverse(r)-Primer aufweisend eine degenerierte Nukleinsäuresequenz gemäß Anspruch 1 kontaktiert,

25 c) eine PCR hiermit durchführt,

 d) aus den erhaltenen Teilsequenzen die Vollängensequenzen der Nukleinsäuren codierend für Proteinsequenzen notwendig zum Aufbau der Aktivität einer Nitrilhydratase generiert und

 e) diese in einen Gastorganismus kloniert und exprimiert.

30 3. Verfahren nach Anspruch 2,

 dadurch gekennzeichnet, dass

 man jeweils Primerpaare aus Primern aufweisend die Nukleinsäuresequenzen A-01f und B-01r oder C-01r sowie B-01f und C-01r bei der PCR einsetzt.

4. Verfahren nach Anspruch 2 und/oder 3, dadurch gekennzeichnet, dass den degenerierten Nukleinsäuresequenzen solche ausgewählt aus der Gruppe bestehend aus:

5	GCCAAGGTCGTC	(Seq. ID NO. 13)
	GGCCGGTCCTG	(Seq. ID NO. 14)
	TCCTTGTACCAGGTC	(Seq. ID NO. 15)
	GCCCGGCC	(Seq. ID NO. 16)
	GGCGCTAATGTTGTT	(Seq. ID NO. 17)
10	TGGCCGGTTCTG	(Seq. ID NO. 18)
	CAAATTCTTTATACCAAGTC	(Seq. ID NO. 19)
	CCATATATCGCATTTCAGCT	(Seq. ID NO. 20)
	GGTCGTGGCCAAG	(Seq. ID NO. 21)
	GGCCGGTCCTG	(Seq. ID NO. 22)
15	TCCTTGTACCAGGTC	(Seq. ID NO. 23)
	GCGCATTTCGGCG	(Seq. ID NO. 24)

vorangestellt sind.

5. Verfahren nach einem oder mehreren der vorhergehenden Ansprüche 2 bis 4,

20	dadurch gekennzeichnet, dass man Primer ausgewählt aus der Gruppe bestehend aus	
	GCCAAGGTCGT <u>Cgcsmrsgcstgg</u>	(Seq. ID NO. 25)
	GGCCGGTCCT <u>Gggscctsccscc</u>	(Seq. ID NO. 26)
	TCCTTGTACCAGGTC <u>ggsgggagscc</u>	(Seq. ID NO. 27)
25	GCCCGCC <u>ggncgcwbsgg</u>	(Seq. ID NO. 28)
	GGCGCTAAAGTTGTT <u>gcnmrrgcntgg</u>	(Seq. ID NO. 29)
	TGGCCGGTTCT <u>Gggnytnccncc</u>	(Seq. ID NO. 30)
	CAAATTCTTTATACCAAGTC <u>ggnggnarncc</u>	(Seq. ID NO. 31)
	CCATATATCGCATTTCAGCT <u>gwngwrtccca</u>	(Seq. ID NO. 32)
30	GGTCGTGGCCAAG <u>gcntggrynga</u>	(Seq. ID NO. 33)
	GGCCGGTCCT <u>Gggnytsccncc</u>	(Seq. ID NO. 34)
	TCCTTGTACCAGGTC <u>ggngggsarncc</u>	(Seq. ID NO. 35)
	GCGCATTTCGGCG <u>swnswrtccca</u>	(Seq. ID NO. 36)

einsetzt.

6. Proteinsequenzen notwendig zum Aufbau der Aktivität einer Nitrilhydratase mit weniger als 100% Homologie auf Aminosäureebene zu derartigen bekannten Proteinsequenzen, wobei die sie codierenden Nukleinsäuresequenzen aus Teilsequenzen generiert werden, die unter stringenten Bedingungen mit den Primern aufweisend die Nukleinsäuresequenzen des Anspruchs 1 ein positives Hybridisierungssignal geben.
5
7. Nukleinsäuresequenz codierend für eine Proteinsequenz nach Anspruch 6.
10
8. Expressionssystem aufweisend eine oder mehrere Nukleinsäuresequenzen nach Anspruch 7.
9. Nitrilhydratase aufweisend Proteinsequenzen für α -Untereinheiten und β -Untereinheiten nach Anspruch 6.
15
10. Verwendung der Nukleinsäuresequenzen nach Anspruch 7 zur Erzeugung von verbesserten Proteinsequenzen notwendig zum Aufbau der Aktivität einer Nitrilhydratase.
11. Verwendung der Nitrilhydratasen nach Anspruch 9 zur Herstellung von organischen Säureamiden und Säuren.
20

Zusammenfassung:

Die vorliegende Erfindung befasst sich mit der Herstellung neuer Nitrilhydratasen. Diese werden vorzugsweise aus nicht kultivierbaren Organismen über eine PCR-basiertes Screening 5 in DNA-Metagenombibliotheken mit speziellen degenerierten Primern erhalten.

Abb. 1

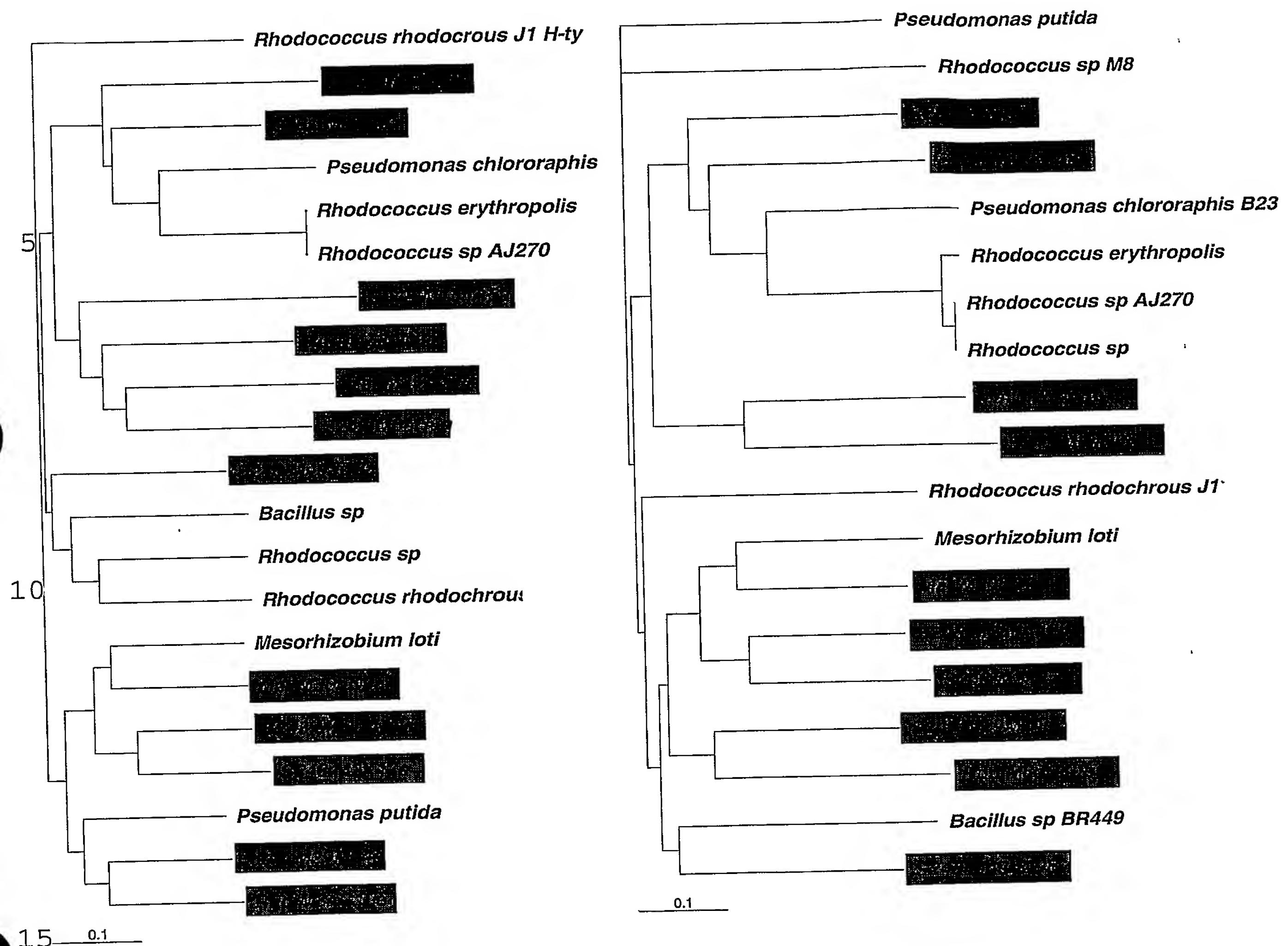


Abb. 2

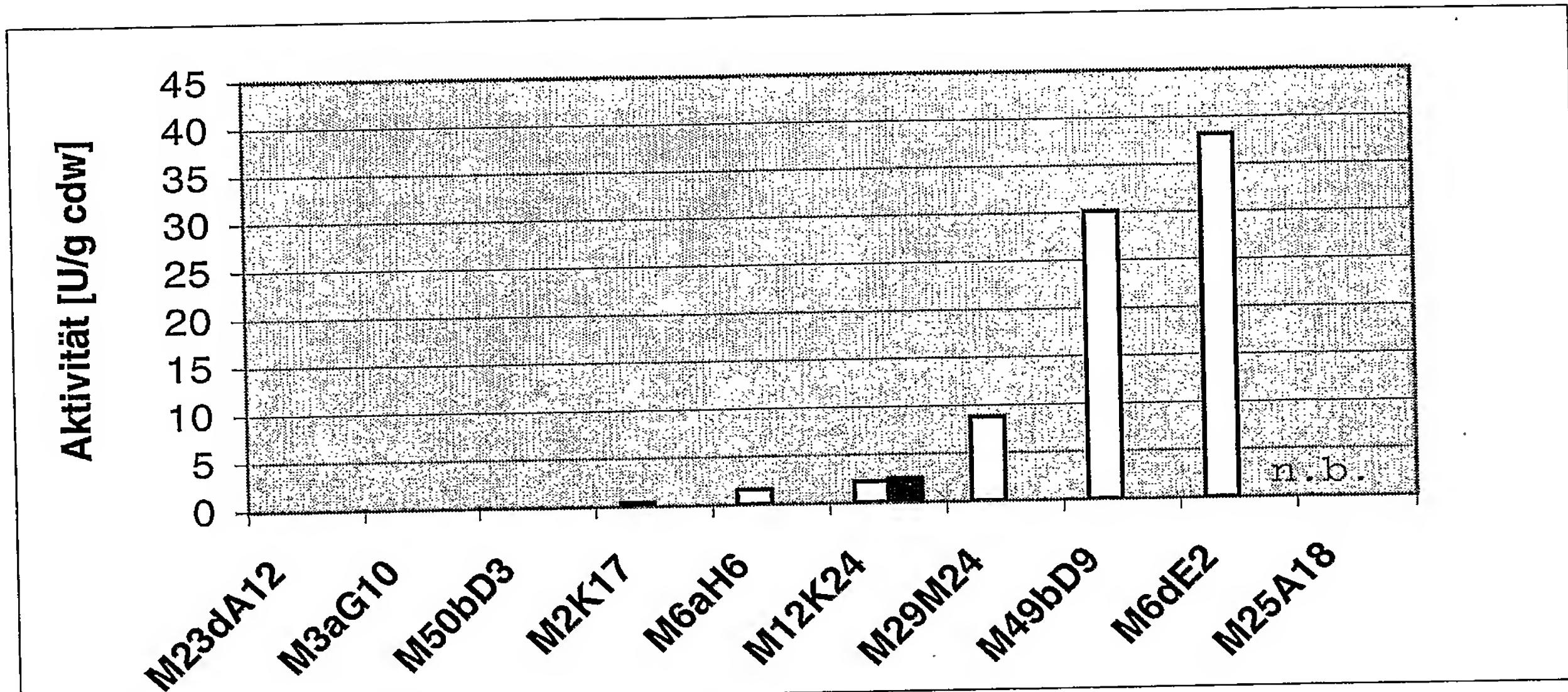


Abb. 3

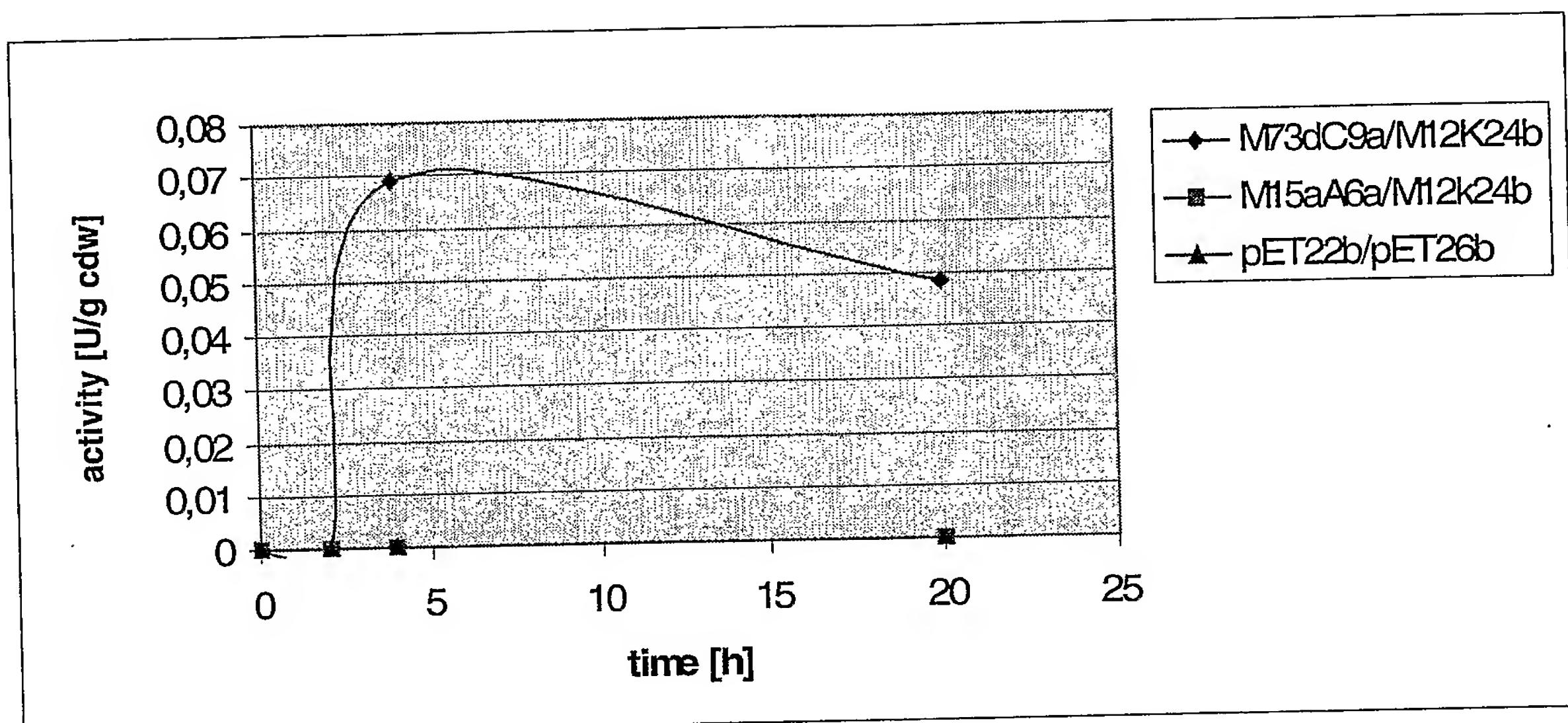


Abb. 4

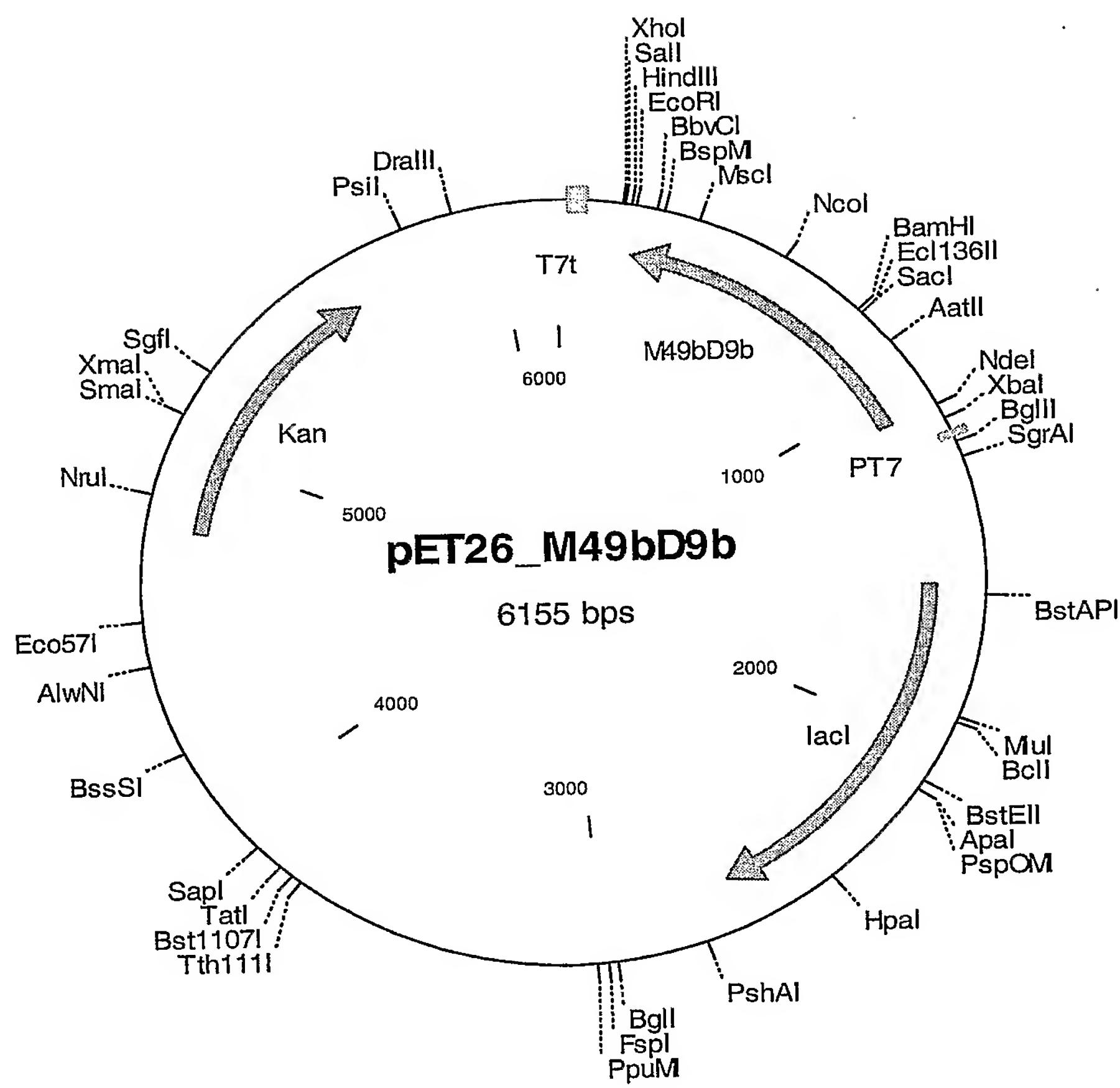


Abb. 5

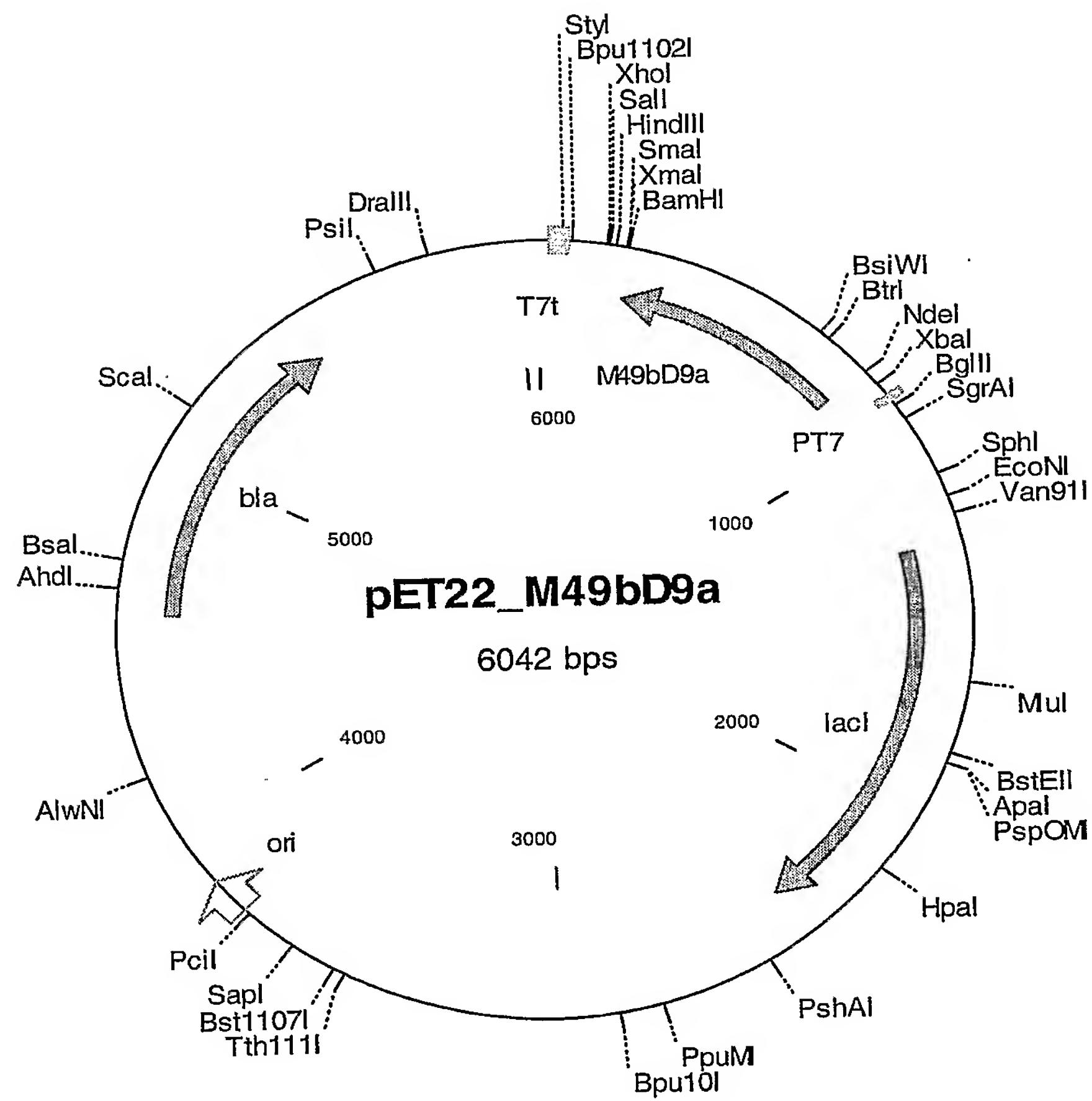


Abb. 6

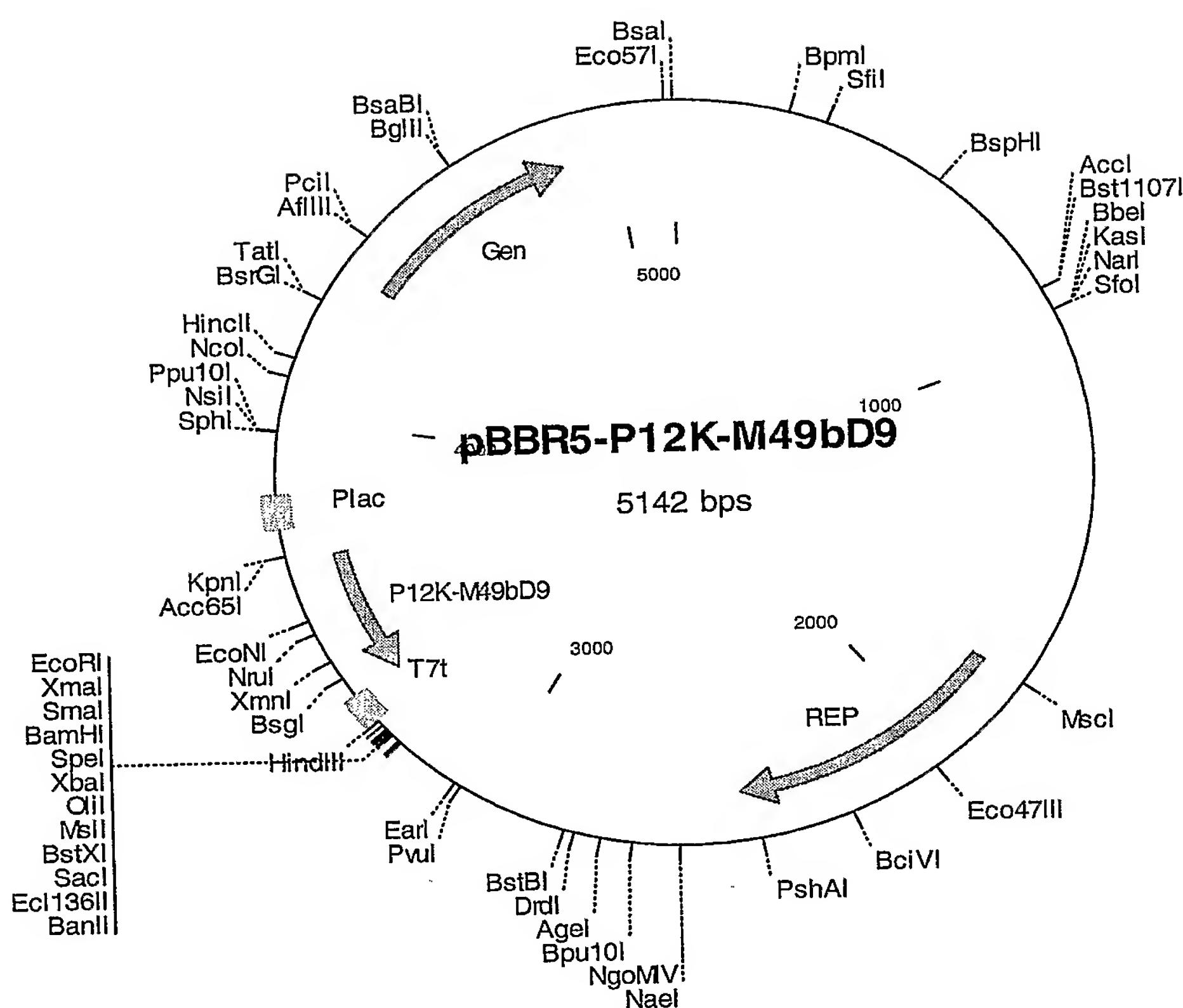
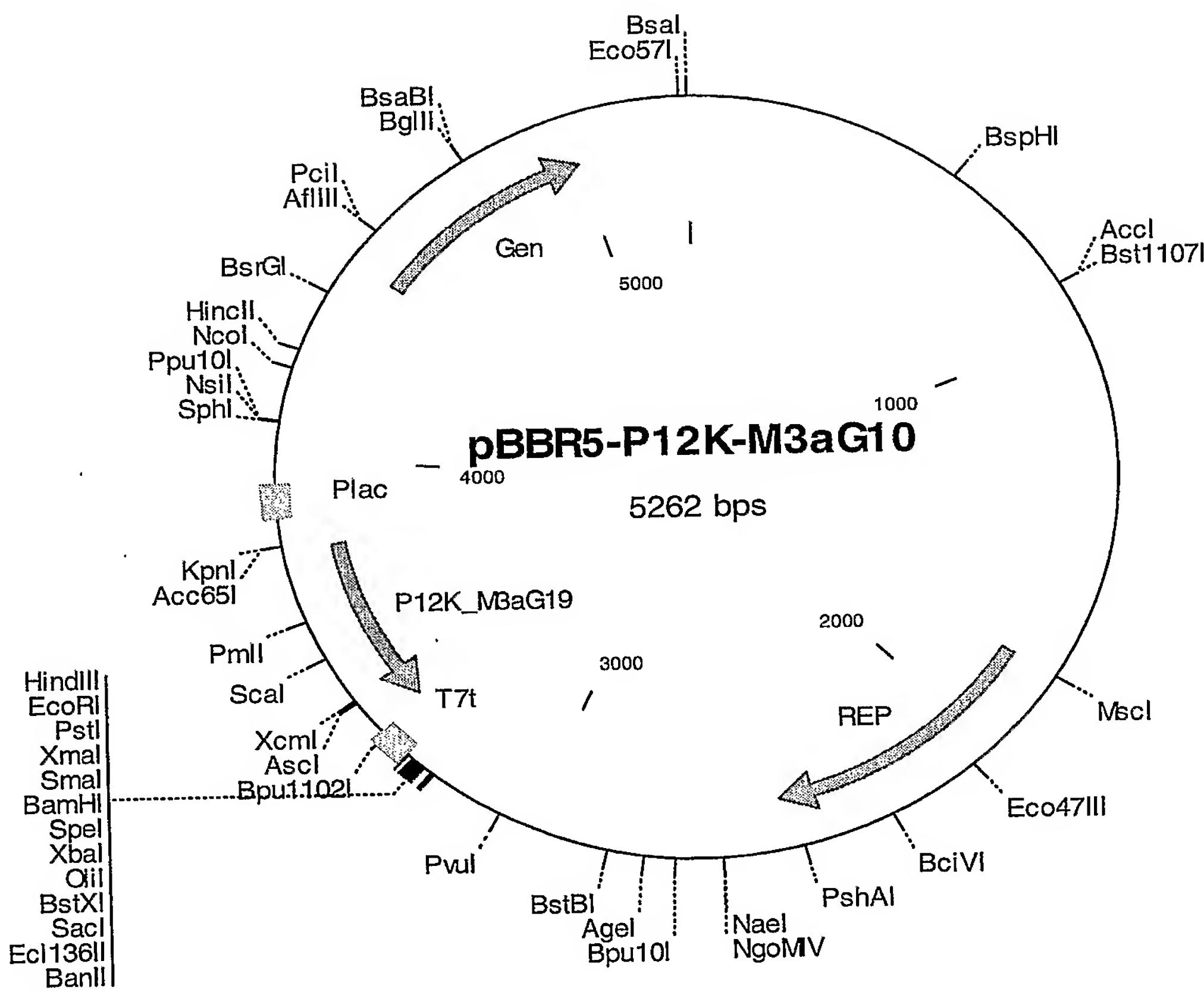


Abb. 7



SEQUENCE LISTING

5 <110> Degussa AG
10 <120> Neue Nitrilhydratasen
15 <130> 040067 AM
20 <160> 86
25 <170> PatentIn version 3.1
30 <210> 1
35 <211> 12
40 <212> DNA
45 <213> Artificial
50 <220>
55 <223> Primer
60 <400> 1
65 gcsmrsgcst gg
70 12
75 <210> 2
80 <211> 11
85 <212> DNA
90 <213> Artificial
95 <220>
100 <223> Primer
105 <400> 2
110 ggsctscsc c
115 11
120 <210> 3
125 <211> 11
130 <212> DNA
135 <213> Artificial
140 <220>
145 <223> Primer
150 <400> 3
155 ggsaggasagsc c
160 11
165 <210> 4
170 <211> 11
175 <212> DNA
180 <213> Artificial
185 <220>
190 <223> Primer

5 <220>
<221> misc_feature
<222> (4)..(4)
<223> a or g or c or t/u

10 <220>
<221> misc_feature
<222> (3)..(3)
<223> a or g or c or t/u

15 <400> 4
ggncgcwbsg g
11

20 <210> 5
<211> 12
<212> DNA
<213> Artificial

25 <220>
<223> Primer

30 <220>
<221> misc_feature
<222> (3)..(3)
<223> a or c

35 <220>
<221> misc_feature
<222> (9)..(9)
<223> a or g or c or t/u

40 <400> 5
gcnmrrgcnt gg
12

45 <210> 6
<211> 11
<212> DNA
<213> Artificial

50 <220>
<223> Primer

55 <220>
<221> misc_feature
<222> (3)..(3)
<223> a or g or c or t/u

60 <220>
<221> misc_feature
<222> (6)..(6)

<223> a or g or c or t/u

5 <220>
 <221> misc_feature
 <222> (9)..(9)
 <223> a or g or c or t/u

10 <400> 6
 ggnnytnccnc c
 11

15 <210> 7
 <211> 11
 <212> DNA
 <213> Artificial

20 <220>
 <223> Primer

25 <220>
 <221> misc_feature
 <222> (3)..(3)
 <223> a or g or c or t/u

30 <220>
 <221> misc_feature
 <222> (6)..(6)
 <223> a or g or c or t/u

35 <220>
 <221> misc_feature
 <222> (9)..(9)
 <223> a or g or c or t/u

40 <400> 7
 ggnggnarnnc c
 11

45 <210> 8
 <211> 11
 <212> DNA
 <213> Artificial

50 <220>
 <223> Primer

55 <220>
 <221> misc_feature
 <222> (3)..(3)
 <223> a or g or c or t/u

<400> 8
gwngwrtccc a
11

5 <210> 9
<211> 11
<212> DNA
<213> Artificial

10 <220>
<223> Primer

15 <220>
<221> misc_feature
<222> (3)..(3)
<223> a or g or c or t/u

20 <220>
<221> misc_feature
<222> (9)..(9)
<223> a or g or c or t/u

25 <400> 9
gcntggryng a
11

30 <210> 10
<211> 11
<212> DNA
<213> Artificial

35 <220>
<223> Primer

40 <220>
<221> misc_feature
<222> (3)..(3)
<223> a or g or c or t/u

45 <220>
<221> misc_feature
<222> (9)..(9)
<223> a or g or c or t/u

50 <400> 10
ggnytsccnnc c
11

55 <210> 11
<211> 11
<212> DNA
<213> Artificial

5 <220>
<221> misc_feature
<222> (3)..(3)
<223> a or g or c or t/u

10 <220>
<221> misc_feature
<222> (9)..(9)
<223> a or g or c or t/u

15 <400> 11
ggngggsarnc c
11

20 <210> 12
<211> 11
<212> DNA
25 <213> Artificial

20 <220>
<223> Primer

30 <220>
<221> misc_feature
<222> (3)..(3)
<223> a or g or c or t/u

35 <400> 12
swnswrtccc a
11

40 <210> 13
<211> 12
<212> DNA
<213> Artificial

45 <220>
<223> Primer

50 <400> 13
gccaaaggctcg tc
12

55 <210> 14
<211> 11
<212> DNA
<213> Artificial

<223> Primer

5 <400> 14
 ggccggtcct g
 11

10 <210> 15
 <211> 15
 <212> DNA
 <213> Artificial

15 <220>
 <223> Primer

20 <400> 15
 tccttgtacc aggtc
 15

25 <210> 16
 <211> 7
 <212> DNA
 <213> Artificial

30 <220>
 <223> Primer

35 <400> 16
 gcccgcc
 7

40 <210> 17
 <211> 15
 <212> DNA
 <213> Artificial

45 <220>
 <223> Primer

50 <400> 17
 ggcgctaaag ttgtt
 15

55 <210> 18
 <211> 12
 <212> DNA
 <213> Artificial

 <220>
 <223> Primer

55 <400> 18
 tggccggttc tg
 12

5 <210> 19
 <211> 20
 <212> DNA
 <213> Artificial
10 <220>
 <223> Primer

15 <400> 19
 caaattcttt ataccaagtc
 20

20 <210> 20
 <211> 20
 <212> DNA
 <213> Artificial

25 <220>
 <223> Primer

30 <400> 20
 ccatatatcg catttcagct
 20
 <210> 21
 <211> 13
 <212> DNA
 <213> Artificial

35 <220>
 <223> Primer

40 <400> 21
 ggtcgtggcc aag
 13

45 <210> 22
 <211> 11
 <212> DNA
 <213> Artificial

50 <220>
 <223> Primer

55 <400> 22
 ggcccggtcct g
 11

 <210> 23
 <211> 15
 <212> DNA
 <213> Artificial

 <220>
 <223> Primer

5 <400> 23
tccttgtacc aggtc
15
10 <210> 24
<211> 13
<212> DNA
<213> Artificial

<220>
<223> Primer

15 <400> 24
gcmcatttcg gcg
13

20 <210> 25
<211> 24
<212> DNA
<213> Artificial

25 <220>
<223> Primer

<400> 25
gccaaaggctcg tcgcsmrsgc stgg
30 24

35 <210> 26
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Primer

40 <400> 26
ggccggtcct gggscctsccs cc
22

45 <210> 27
<211> 26
<212> DNA
<213> Artificial

50 <220>
<223> Primer

<400> 27
55 tccttgtacc aggtcggsgg sagscc
26

<210> 28

<211> 18
<212> DNA
<213> Artificial

5 <220>
<223> Primer

<220>
<221> misc_feature
10 <222> (10)..(10)
<223> a or c or g or t/u

<400> 28
15 gccccggcggc ggcwbsgg
18

20 <210> 29
<211> 27
<212> DNA
<213> Artificial

25 <220>
<223> Primer

<220>
<221> misc_feature
<222> (18)..(18)
30 <223> a or c or g or t/u

<220>
<221> misc_feature
35 <222> (24)..(24)
<223> a or c or g or t/u

40 <400> 29
45 ggcgctaaag ttgttgcnmr rgcntgg
27

45 <210> 30
<211> 22
<212> DNA
<213> Artificial

50 <220>
<223> Primer

<220>
<221> misc_feature
<222> (15)..(15)
55 <223> a or c or g or t/u

<220>
<221> misc_feature

<222> (18)..(18)
<223> a or c or g or t/u

5 <220>
<221> misc_feature
<222> (21)..(21)
<223> a or c or g or t/u

10 <400> 30
tggccgggttc tgggnytncc nc
22

15 <210> 31
<211> 31
<212> DNA
<213> Artificial

20 <220>
<223> Primer

25 <220>
<221> misc_feature
<222> (23)..(23)
<223> a or c or g or t/u

30 <220>
<221> misc_feature
<222> (26)..(26)
<223> a or c or g or t/u

35 <220>
<221> misc_feature
<222> (29)..(29)
<223> a or c or g or t/u

40 <400> 31
caaattcttt ataccaagtc ggnggnarnnc c
31

45 <210> 32
<211> 31
<212> DNA
<213> Artificial

<220>
<223> Primer

55 <220>
<221> misc_feature
<222> (23)..(23)
<223> a or c or g or t/u

5 <400> 32
ccatatatcg catttcagct gwngwrtccc a
31
10 <210> 33
<211> 24
<212> DNA
<213> Artificial
15 <220>
<221> misc_feature
<222> (16)..(16)
<223> a or c or g or t/u
20 <220>
<221> misc_feature
<222> (22)..(22)
<223> a or c or g or t/u
25
30 <400> 33
ggtcgtggcc aaggcntggr yngaa
24
35 <210> 34
<211> 22
<212> DNA
<213> Artificial
40 <220>
<221> misc_feature
<222> (14)..(14)
<223> a or c or g or t/u
45 <220>
<221> misc_feature
<222> (20)..(20)
<223> a or c or g or t/u
50
55 <400> 34
ggccgggtcct gggnytscnn cc
22
<210> 35
<211> 26
<212> DNA

```
<213> Artificial

<220>
<223> Primer
5

<220>
<221> misc_feature
<222> (18)..(18)
<223> a or c or g or t/u
10

<220>
<221> misc_feature
<222> (24)..(24)
15 <223> a or c or g or t/u

<400> 35
tccttgtacc aggtcggnngg sarncc
20 26

<210> 36
<211> 24
25 <212> DNA
<213> Artificial

<220>
<223> Primer
30

<220>
<221> misc_feature
<222> (16)..(16)
<223> a or c or g or t/u
35

<400> 36
gcgcatttcg gcgswnswrt cccaa
40 24

<210> 37
<211> 705
<212> DNA
45 <213> Unknown

<220>
<223> Metagenome - alpha unit nitrile hydratase - M6aH6
50

<220>
<221> CDS
<222> (1)..(705)
<223>

55 <400> 37
atg agc aag cac gta cac gat tac tac gcg aag aag aag cac gac cac
48
Met Ser Lys His Val His Asp Tyr Tyr Ala Lys Lys Lys His Asp His
15
1 5 10 15
```

gat cat gac cac gac gtc cac gag gcg atc gag gac cg^g gac gag ggt
 96
 Asp His Asp His Asp Val His Glu Ala Ile Glu Asp Arg Asp Glu Gly
 5 20 25 30
 ccg ccg tcg gaa ttc gag atc atg agc cgc gcc atg cag gag ctg ctg
 144
 Pro Pro Ser Glu Phe Glu Ile Met Ser Arg Ala Met Gln Glu Leu Leu
 10 35 40 45
 gaa gag aag ggc gtc gtc acc gcg gag cag gtc cg^g cgc agc atg gag
 192
 Glu Glu Lys Gly Val Val Thr Ala Glu Gln Val Arg Arg Ser Met Glu
 15 50 55 60
 aag ttc gag gaa gag ctg ccc tac cg^g ggg gcg cg^g gtc gtc g^g cac
 240
 Lys Phe Glu Glu Leu Pro Tyr Arg Gly Ala Arg Val Val Ala His
 0 65 70 75 80
 gcc tgg acc gac ccg gaa ttc aag aag cg^g ctg ctg g^g gac ggc aag
 288
 Ala Trp Thr Asp Pro Glu Phe Lys Lys Arg Leu Leu Ala Asp Gly Lys
 25 85 90 95
 gcc gtc tcg gag ttc ggc atc gat ttg gaa gcc gag cg^g ctg atc
 336
 Ala Ala Val Ser Glu Phe Gly Ile Asp Leu Glu Ala Glu Arg Leu Ile
 30 100 105 110
 gcg gtc gc^g aac aca acg gac gtg cac aac gtc atc gtg tgc acg ctg
 384
 Ala Val Ala Asn Thr Thr Asp Val His Asn Val Ile Val Cys Thr Leu
 35 115 120 125
 tgc tcg tgc tac ccg cg^g acg ctg ctc ggc atg ccg ccg acc tgg tac
 432
 Cys Ser Cys Tyr Pro Arg Thr Leu Leu Gly Met Pro Pro Thr Trp Tyr
 40 130 135 140
 aag agc gat aac tac cg^g tcg cg^g gtg gtc tac gaa ccg cg^g g^g gtg
 480
 Lys Ser Asp Asn Tyr Arg Ser Arg Val Val Tyr Glu Pro Arg Ala Val
 45 145 150 155 160
 ctg aag gaa ttc ggc acc gtg ctg ccg gag cg^g gtc acc gtg cg^g gtg
 528
 Leu Lys Glu Phe Gly Thr Val Leu Pro Glu Arg Val Thr Val Arg Val
 50 165 170 175
 cac gac tcc aac gcc gac atg cg^g tac gtg gtg atc ccc atg cg^g ccg
 576
 His Asp Ser Asn Ala Asp Met Arg Tyr Val Val Ile Pro Met Arg Pro
 55 180 185 190
 cag gga acc gag ggc tgg agc gag gag cg^g ctc g^g gag ctg ctg acg
 624
 Gln Gly Thr Glu Gly Trp Ser Glu Glu Arg Leu Ala Glu Leu Leu Thr

195 200 205

cgc gac acg ctg gtg ggg gtc acc gtg cca aaa gtg gaa gtc gga agt
672 672
5 Arg Asp Thr Leu Val Gly Val Thr Val Pro Lys Val Glu Val Gly Ser
210 215 220

cga aag tcg aaa ggc gga agt aaa acc cgc tag
705 705
10 Arg Lys Ser Lys Gly Gly Ser Lys Thr Arg
225 230

15 <210> 38
<211> 234
<212> PRT
<213> Unknown

20 <220>
<223> Metagenome - alpha unit nitrile hydratase - M6aH6

<400> 38

25 Met Ser Lys His Val His Asp Tyr Tyr Ala Lys Lys Lys His Asp His
1 5 10 15

Asp His Asp His Asp Val His Glu Ala Ile Glu Asp Arg Asp Glu Gly
20 25 30

30 Pro Pro Ser Glu Phe Glu Ile Met Ser Arg Ala Met Gln Glu Leu Leu
35 40 45

35 Glu Glu Lys Gly Val Val Thr Ala Glu Gln Val Arg Arg Ser Met Glu
50 55 60

40 Lys Phe Glu Glu Glu Leu Pro Tyr Arg Gly Ala Arg Val Val Ala His
65 70 75 80

45 Ala Trp Thr Asp Pro Glu Phe Lys Lys Arg Leu Leu Ala Asp Gly Lys
85 90 95

50 Ala Ala Val Ser Glu Phe Gly Ile Asp Leu Glu Ala Glu Arg Leu Ile
100 105 110

Ala Val Ala Asn Thr Thr Asp Val His Asn Val Ile Val Cys Thr Leu
115 120 125

55 Cys Ser Cys Tyr Pro Arg Thr Leu Leu Gly Met Pro Pro Thr Trp Tyr
130 135 140

Lys Ser Asp Asn Tyr Arg Ser Arg Val Val Tyr Glu Pro Arg Ala Val
145 150 155 160

5 Leu Lys Glu Phe Gly Thr Val Leu Pro Glu Arg Val Thr Val Arg Val
165 170 175

10 His Asp Ser Asn Ala Asp Met Arg Tyr Val Val Ile Pro Met Arg Pro
180 185 190

15 Gln Gly Thr Glu Gly Trp Ser Glu Glu Arg Leu Ala Glu Leu Leu Thr
195 200 205

20 Arg Asp Thr Leu Val Gly Val Thr Val Pro Lys Val Glu Val Gly Ser
210 215 220

25 Arg Lys Ser Lys Gly Gly Ser Lys Thr Arg
225 230

30 <210> 39
<211> 657
<212> DNA
<213> Unknown

35 <220>
<223> Metagenome - alpha unit nitrile hydratase - M12K24

40 <220>
<221> CDS
<222> (1)..(657)
<223>

45 <400> 39
atg tcg gac gga aca aca att gga atc cag gcc gcg aca acc ctt cga
48 Met Ser Asp Gly Thr Thr Ile Gly Ile Gln Ala Ala Thr Thr Leu Arg
1 5 10 15

50 tca gcc atg aac att cca gct cgt gaa ttc gcc ctc cag cgc act gcg
96 Ser Ala Met Asn Ile Pro Ala Arg Glu Phe Ala Leu Gln Arg Thr Ala
20 25 30

55 ccg gtc gag cag cgt gtc gac gcg atc cag gcg ctc gac gaa cgc
144 Pro Val Glu Gln Arg Val Asp Ala Ile Gln Ala Ala Leu Asp Glu Arg
35 40 45

59 ggt ttg aac gcc agt gac gca gtc cag gaa ttg agc cac ctg gcg gag
192 Gly Leu Asn Ala Ser Asp Ala Val Gln Glu Leu Ser His Leu Ala Glu
50 55 60

gag caa tgg att ccg cgc aat ggc gcg cggt gtc gtc gcc aaa gcc tgg
 240
 Glu Gln Trp Ile Pro Arg Asn Gly Ala Arg Val Val Ala Lys Ala Trp
 65 70 75 80
 5 gtc gac ccg gaa ttc cgc gcg cggt ctt ctg gcc gac ggt cgc gcc gcc
 288
 Val Asp Pro Glu Phe Arg Ala Arg Leu Leu Ala Asp Gly Arg Ala Ala
 85 90 95
 10 gtt gcc gaa ctg ggc ctc tcg atg ccg aag cat cac ccg cac ctc gtg
 336
 Val Ala Glu Leu Gly Leu Ser Met Pro Lys His His Arg His Leu Val
 100 105 110
 15 gtg ctg gag aac acg ccg agc gtg cag aac gtc atc tgc tgc acg cag
 384
 Val Leu Glu Asn Thr Pro Ser Val Gln Asn Val Ile Cys Cys Thr Gln
 115 120 125
 20 tgc tcg tgc acg gcg ttc acg atc atc gga ctg ccg ccc gac tgg tac
 432
 Cys Ser Cys Thr Ala Phe Thr Ile Ile Gly Leu Pro Pro Asp Trp Tyr
 130 135 140
 25 aag gac ctg gaa tac cgc gcg cga gtc gtc ccg gag tcg cgc acc gtg
 480
 Lys Asp Leu Glu Tyr Arg Ala Arg Val Val Arg Glu Ser Arg Thr Val
 145 150 155 160
 30 ctg aag gag atg gga ctg gat ctg cct ccg gat gtc gaa att cgc gtc
 528
 Leu Lys Glu Met Gly Leu Asp Leu Pro Arg Asp Val Glu Ile Arg Val
 165 170 175
 35 tgg gat acc act gcc gac acg cgc tac atg gta ttg ccg gta cag ccg
 576
 Trp Asp Thr Thr Ala Asp Thr Arg Tyr Met Val Leu Pro Val Gln Pro
 180 185 190
 40 ccg gaa acc atc ggc tgg ccc gag gag aaa ctg gtc gac atc gtg acg
 624
 Pro Glu Thr Ile Gly Trp Pro Glu Glu Lys Leu Val Asp Ile Val Thr
 195 200 205
 45 cgc gac ggc atg atc ggc gtc gcg ccg gta tag
 657
 Arg Asp Gly Met Ile Gly Val Ala Arg Val
 210 215
 50
 55 <210> 40
 <211> 218
 <212> PRT
 <213> Unknown
 <220>
 <223> Metagenome - alpha unit nitrile hydratase - M12K24

<400> 40

Met Ser Asp Gly Thr Thr Ile Gly Ile Gln Ala Ala Thr Thr Leu Arg
1 5 10 15

5

Ser Ala Met Asn Ile Pro Ala Arg Glu Phe Ala Leu Gln Arg Thr Ala
20 25 30

10

Pro Val Glu Gln Arg Val Asp Ala Ile Gln Ala Ala Leu Asp Glu Arg
35 40 45

15

Gly Leu Asn Ala Ser Asp Ala Val Gln Glu Leu Ser His Leu Ala Glu
50 55 60

20

Glu Gln Trp Ile Pro Arg Asn Gly Ala Arg Val Val Ala Lys Ala Trp
65 70 75 80

25

Val Asp Pro Glu Phe Arg Ala Arg Leu Leu Ala Asp Gly Arg Ala Ala
85 90 95

30

Val Ala Glu Leu Gly Leu Ser Met Pro Lys His His Arg His Leu Val
100 105 110

35

Val Leu Glu Asn Thr Pro Ser Val Gln Asn Val Ile Cys Cys Thr Gln
115 120 125

40

Cys Ser Cys Thr Ala Phe Thr Ile Ile Gly Leu Pro Pro Asp Trp Tyr
130 135 140

45

Lys Asp Leu Glu Tyr Arg Ala Arg Val Val Arg Glu Ser Arg Thr Val
145 150 155 160Leu Lys Glu Met Gly Leu Asp Leu Pro Arg Asp Val Glu Ile Arg Val
165 170 175

50

Trp Asp Thr Thr Ala Asp Thr Arg Tyr Met Val Leu Pro Val Gln Pro
180 185 190Pro Glu Thr Ile Gly Trp Pro Glu Glu Lys Leu Val Asp Ile Val Thr
195 200 205

55

Arg Asp Gly Met Ile Gly Val Ala Arg Val
210 215

<211> 699
<212> DNA
<213> Unknown

5 <220>
<223> Metagenome - alpha unit nitrile hydratase - M29M24

<220>
<221> CDS
10 <222> (1)..(699)
<223>

<400> 41
atg gag acg aca cgc cgc ggc ttc ttg aag aaa gcc ggg acc gcc gcc
15 48
Met Glu Thr Thr Arg Arg Gly Phe Leu Lys Lys Ala Gly Thr Ala Ala
1 5 10 15
ggt gcg acg gct gca ccc gtt gga ctc gcc aag atc gct cac tcc cac
20 96
Gly Ala Thr Ala Ala Pro Val Gly Leu Ala Lys Ile Ala His Ser His
20 25 30
gag cac cag gcc gtt cct tcc gac ctc acg ctc cgg gtc aag tcc ctc
25 144
Glu His Gln Ala Val Pro Ser Asp Leu Thr Leu Arg Val Lys Ser Leu
35 35 40 45
gaa tcg ctg ctg gtc gag aag ggt ctc gtg gac cgt gag gcc ctc gac
30 192
Glu Ser Leu Leu Val Glu Lys Gly Leu Val Asp Arg Glu Ala Leu Asp
50 50 55 60
gtg ctc gtc gat acc tac gag aac aag atc ggt ccg cga aac ggc gct
35 240
Val Leu Val Asp Thr Tyr Glu Asn Lys Ile Gly Pro Arg Asn Gly Ala
65 65 70 75 80
cgc gtc gtc gcg cgg gcg tgg gtc gat ccc gcg tac aaa gag cgc ctg
40 288
Arg Val Val Ala Arg Ala Trp Val Asp Pro Ala Tyr Lys Glu Arg Leu
85 85 90 95
ctg aaa gac gcc acc tcg gcg atc gcc gag ctc ggt tac acc gga gcc
45 336
Leu Lys Asp Ala Thr Ser Ala Ile Ala Glu Leu Gly Tyr Thr Gly Ala
100 100 105 110
cag ggt gag cac atg gtg gcg ctc gag aat acc ccc gcg gtg cac aac
50 384
Gln Gly Glu His Met Val Ala Leu Glu Asn Thr Pro Ala Val His Asn
115 115 120 125
ctc gtc gtt tgc acg ctc tgc tat cca tgg cgg gtg ctc ggt
55 432
Leu Val Val Cys Thr Leu Cys Ser Cys Tyr Pro Trp Pro Val Leu Gly
130 135 140

ctg ccc ccg gtc tgg tac aaa tcg gcg ccc tac cga tcg cgc tcg gtc
 480
 Leu Pro Pro Val Trp Tyr Lys Ser Ala Pro Tyr Arg Ser Arg Ser Val
 145 150 155 160
 5 atc gat ccg cgc ggc gtt ctc ggc gag ttc ggg ctc gag ctg ccg gaa
 528
 Ile Asp Pro Arg Gly Val Leu Gly Glu Phe Gly Leu Glu Leu Pro Glu
 165 170 175
 10 ggg gtc gag gtg cgc gtc tgg gac tcg acg gcg gag ctc cgg tat ctc
 576
 Gly Val Glu Val Arg Val Trp Asp Ser Thr Ala Glu Leu Arg Tyr Leu
 180 185 190
 15 gtt ttg ccg gag cgg ccc gaa ggc acg gcg caa ctg agc gaa gaa gcg
 624
 Val Leu Pro Glu Arg Pro Glu Gly Thr Ala Gln Leu Ser Glu Glu Ala
 195 200 205
 20 ctc gcg gat ctc acc cgg gat gcc atg atc ggc gtc gcg aaa gtc
 672
 Leu Ala Asp Leu Val Thr Arg Asp Ala Met Ile Gly Val Ala Lys Val
 210 215 220
 25 tcg ttg ccc gcg ggc ggc gaa tga
 699
 Ser Leu Pro Ala Gly Gly Ala Glu
 225 230
 30
 <210> 42
 <211> 232
 <212> PRT
 35 <213> Unknown
 <220>
 <223> Metagenome - alpha unit nitrile hydratase - M29M24
 40 <400> 42
 Met Glu Thr Thr Arg Arg Gly Phe Leu Lys Lys Ala Gly Thr Ala Ala
 1 5 10 15
 45 Gly Ala Thr Ala Ala Pro Val Gly Leu Ala Lys Ile Ala His Ser His
 20 25 30
 50 Glu His Gln Ala Val Pro Ser Asp Leu Thr Leu Arg Val Lys Ser Leu
 35 40 45
 55 Glu Ser Leu Leu Val Glu Lys Gly Leu Val Asp Arg Glu Ala Leu Asp
 50 55 60
 Val Leu Val Asp Thr Tyr Glu Asn Lys Ile Gly Pro Arg Asn Gly Ala
 65 70 75 80

Arg Val Val Ala Arg Ala Trp Val Asp Pro Ala Tyr Lys Glu Arg Leu
85 90 95

5 Leu Lys Asp Ala Thr Ser Ala Ile Ala Glu Leu Gly Tyr Thr Gly Ala
100 105 110

10 Gln Gly Glu His Met Val Ala Leu Glu Asn Thr Pro Ala Val His Asn
115 120 125

15 Leu Val Val Cys Thr Leu Cys Ser Cys Tyr Pro Trp Pro Val Leu Gly
130 135 140

20 Leu Pro Pro Val Trp Tyr Lys Ser Ala Pro Tyr Arg Ser Arg Ser Val
145 150 155 160

Ile Asp Pro Arg Gly Val Leu Gly Glu Phe Gly Leu Glu Leu Pro Glu
165 170 175

25 Gly Val Glu Val Arg Val Trp Asp Ser Thr Ala Glu Leu Arg Tyr Leu
180 185 190

30 Val Leu Pro Glu Arg Pro Glu Gly Thr Ala Gln Leu Ser Glu Glu Ala
195 200 205

35 Leu Ala Asp Leu Val Thr Arg Asp Ala Met Ile Gly Val Ala Lys Val
210 215 220

40 Ser Leu Pro Ala Gly Gly Ala Glu
225 230

45 <210> 43
<211> 639
<212> DNA
<213> Unknown

50 <220>
<223> Metagenome - alpha unit nitrile hydratase - M2K17

55 <220>
<221> CDS
<222> (1)..(639)
<223>
<400> 43
atg cct gac gac cat gcc cat ccg gat gat cat gcg cat ggc tcg gaa
48
Met Pro Asp Asp His Ala His Pro Asp Asp His Ala His Gly Ser Glu

1	5	10	15
ttg tcc gag atg gat atc cgg gtg cg ^g ctg gag acc atc ctg acc			
96 Leu Ser Glu Met Asp Ile Arg Val Arg Ala Leu Glu Thr Ile Leu Thr			
5 20 25 30			
gag aag ggc tat gtc gat ccg gc ^g ctc gac cg ^g atc gtc gag gc ^g			
144 Glu Lys Gly Tyr Val Asp Pro Ala Ala Leu Asp Arg Ile Val Glu Ala			
10 35 40 45			
ttc gag acc agg atc ggc ccg cat atc ggc gcc cgt atc gtg gca cg ^g			
192 Phe Glu Thr Arg Ile Gly Pro His Ile Gly Ala Arg Ile Val Ala Arg			
15 50 55 60			
gct tgg gcc gac gcc gaa ttc aag cg ^g cg ^g ctg ctc gcc gac gc ^g acc			
240 Ala Trp Ala Asp Ala Glu Phe Lys Arg Arg Leu Leu Ala Asp Ala Thr			
20 65 70 75 80			
gag gc ^g gc ^g aat tcg ctg ggt cat gc ^g agc ccg gtc ggc agc cat ctg			
288 Glu Ala Ala Asn Ser Leu Gly His Ala Ser Pro Val Gly Ser His Leu			
25 85 90 95			
atc gc ^g gtc gag aac acg ccg cag acc cac aac ctc gtc gtc tgc act			
336 Ile Ala Val Glu Asn Thr Pro Gln Thr His Asn Leu Val Val Cys Thr			
30 100 105 110			
ttg tgc tcg tgt tat ccg tgg gag gtg ctg gga ttg ccg ccg gtc tgg			
384 Leu Cys Ser Cys Tyr Pro Trp Glu Val Leu Gly Leu Pro Pro Val Trp			
35 115 120 125			
tac aaa tcc gct gcc tac cg ^g tcg cg ^g gtg gtg atc gac ccc aag gg ^g			
432 Tyr Lys Ser Ala Ala Tyr Arg Ser Arg Val Val Ile Asp Pro Lys Gly			
40 130 135 140			
gtc ctc gcc gag ttc ggc ctg acc ctg cca ccg gag acc ggg atc cg ^g			
480 Val Leu Ala Glu Phe Gly Leu Thr Leu Pro Pro Glu Thr Gly Ile Arg			
45 145 150 155 160			
atc tgg gat tcg acc gcc gag acc cg ^g ttt ctg gtg gtg ccg atg cg ^g			
528 Ile Trp Asp Ser Thr Ala Glu Thr Arg Phe Leu Val Val Pro Met Arg			
50 165 170 175			
ccc ccc ggc acc gca ggc tgg agc gag gaa cg ^g ctc gcc gaa ctc gtc			
576 Pro Pro Gly Thr Ala Gly Trp Ser Glu Glu Arg Leu Ala Glu Leu Val			
55 180 185 190			
acc cg ^g gac agc atg atc ggc act ggt ctg gcc ggg gcg ccg cag gag			
624			

Thr Arg Asp Ser Met Ile Gly Thr Gly Leu Ala Gly Ala Pro Gln Glu
195 200 205

5 atg gcc tcg gca tga
639
Met Ala Ser Ala
210

10 <210> 44
<211> 212
<212> PRT
<213> Unknown

15 <220>
<223> Metagenome - alpha unit nitrile hydratase - M2K17

<400> 44

20 Met Pro Asp Asp His Ala His Pro Asp Asp His Ala His Gly Ser Glu
1 5 10 15

25 Leu Ser Glu Met Asp Ile Arg Val Arg Ala Leu Glu Thr Ile Leu Thr
20 25 30

30 Glu Lys Gly Tyr Val Asp Pro Ala Ala Leu Asp Arg Ile Val Glu Ala
35 40 45

35 Phe Glu Thr Arg Ile Gly Pro His Ile Gly Ala Arg Ile Val Ala Arg
50 55 60

40 Ala Trp Ala Asp Ala Glu Phe Lys Arg Arg Leu Leu Ala Asp Ala Thr
65 70 75 80

45 Glu Ala Ala Asn Ser Leu Gly His Ala Ser Pro Val Gly Ser His Leu
85 90 95

50 Ile Ala Val Glu Asn Thr Pro Gln Thr His Asn Leu Val Val Cys Thr
100 105 110

55 Leu Cys Ser Cys Tyr Pro Trp Glu Val Leu Gly Leu Pro Pro Val Trp
115 120 125

55 Tyr Lys Ser Ala Ala Tyr Arg Ser Arg Val Val Ile Asp Pro Lys Gly
130 135 140

55 Val Leu Ala Glu Phe Gly Leu Thr Leu Pro Pro Glu Thr Gly Ile Arg
145 150 155 160

Ile Trp Asp Ser Thr Ala Glu Thr Arg Phe Leu Val Val Pro Met Arg
 165 170 175

5 Pro Pro Gly Thr Ala Gly Trp Ser Glu Glu Arg Leu Ala Glu Leu Val
 180 185 190

10 Thr Arg Asp Ser Met Ile Gly Thr Gly Leu Ala Gly Ala Pro Gln Glu
 195 200 205

15 Met Ala Ser Ala
 210

20 <210> 45

<211> 696

<212> DNA

<213> Unknown

25 <220>

<221> CDS

<222> (1)..(696)

<223>

30 <400> 45

atg cgt tcg ccc ggt gag gcc tca gca acg caa cca gcg ctc att cgg
 48

Met Arg Ser Pro Gly Glu Ala Ser Ala Thr Gln Pro Ala Leu Ile Arg
 1 5 10 15

35

ctg cat gat cga gct ggc ggc gtt cga tca ttg cgc ggc aaa agg tct
 96

Leu His Asp Arg Ala Gly Gly Val Arg Ser Leu Arg Gly Lys Arg Ser
 20 25 30

40

cat cgc gcc gga tcg cat cct cgg ggc gct cgc gca tcc gtc gcc aca
 144

His Arg Ala Gly Ser His Pro Arg Gly Ala Arg Ala Ser Val Ala Thr
 35 40 45

45

ggg tgg ttc gtt ccg ttc tcg gcc agg ctc gcc cgg aaa ggc atc gct
 192

Gly Trp Phe Val Pro Phe Ser Ala Arg Leu Ala Arg Lys Gly Ile Ala
 50 55 60

50

cct ccg gcc gag atc gcc gag cgg atc gcc gtc acc gat cgc gca tca
 240

Pro Pro Ala Glu Ile Ala Glu Arg Ile Ala Val Thr Asp Arg Ala Ser
 65 70 75 80

55

ccg gca atg ggc gct cgc atg gtc gcc aag gcc tgg acc gat ccc gcc
 288

Pro Ala Met Gly Ala Arg Met Val Ala Lys Ala Trp Thr Asp Pro Ala
 85 90 95

ttc cgc acc ctg ctc ttg gaa gac gga acc cgc gcg gcg gaa tcg ctc
 336
 Phe Arg Thr Leu Leu Leu Glu Asp Gly Thr Arg Ala Ala Glu Ser Leu
 5 100 105 110

 ggc atc atg atg cgc ggc gcc ccg cct ctc ggt gtg ctg gag aat acg
 384
 Gly Ile Met Met Arg Gly Ala Pro Pro Leu Gly Val Leu Glu Asn Thr
 10 115 120 125

 ccc gag att cat cac ctc gtc gtt tgc acg ctg tgc agt tgt tac ccg
 432
 Pro Glu Ile His His Leu Val Val Cys Thr Leu Cys Ser Cys Tyr Pro
 15 130 135 140

 cgc gcg gtg ctg ggc tat ccg ccg ttc tgg ttc aaa tcc gcc gcc tac
 480
 Arg Ala Val Leu Gly Tyr Pro Pro Phe Trp Phe Lys Ser Ala Ala Tyr
 20 145 150 155 160

 cg^g gca cgt gcg gtg cgc gac ccg cgc ggt ctg atc gcc gaa tgg ggc
 528
 Arg Ala Arg Ala Val Arg Asp Pro Arg Gly Leu Ile Ala Glu Trp Gly
 25 165 170 175

 acc atg ctg ccc gac gat gtc cgc gtg cga gtg gtg gac agt acg gcc
 576
 Thr Met Leu Pro Asp Asp Val Arg Val Arg Val Val Asp Ser Thr Ala
 30 180 185 190

 gac tat cgc tgg atg gtt ctg ccg gtg cgg ccg gcc ggc act gcg ggc
 624
 Asp Tyr Arg Trp Met Val Leu Pro Val Arg Pro Ala Gly Thr Ala Gly
 35 195 200 205

 tgg gat gag gag cgc ctc gcc gca atc gta cgc gaa ggc gat atg atc
 672
 Trp Asp Glu Glu Arg Leu Ala Ala Ile Val Arg Glu Gly Asp Met Ile
 40 210 215 220

 ggg gtg acc atc cct cgt ctt taa
 696
 Gly Val Thr Ile Pro Arg Leu
 45 225 230

 <210> 46
 <211> 231
 50 <212> PRT
 <213> Unknown

 <220>
 <223> Metagenome - alpha unit nitrile hydratase - M15aA6
 55
 <400> 46

 Met Arg Ser Pro Gly Glu Ala Ser Ala Thr Gln Pro Ala Leu Ile Arg
 1 5 10 15

Leu His Asp Arg Ala Gly Gly Val Arg Ser Leu Arg Gly Lys Arg Ser
20 25 30

5 His Arg Ala Gly Ser His Pro Arg Gly Ala Arg Ala Ser Val Ala Thr
35 40 45

10 Gly Trp Phe Val Pro Phe Ser Ala Arg Leu Ala Arg Lys Gly Ile Ala
50 55 60

15 Pro Pro Ala Glu Ile Ala Glu Arg Ile Ala Val Thr Asp Arg Ala Ser
65 70 75 80

20 Pro Ala Met Gly Ala Arg Met Val Ala Lys Ala Trp Thr Asp Pro Ala
85 90 95

25 Phe Arg Thr Leu Leu Leu Glu Asp Gly Thr Arg Ala Ala Glu Ser Leu
100 105 110

30 Gly Ile Met Met Arg Gly Ala Pro Pro Leu Gly Val Leu Glu Asn Thr
115 120 125

35 Pro Glu Ile His His Leu Val Val Cys Thr Leu Cys Ser Cys Tyr Pro
130 135 140

40 Arg Ala Val Leu Gly Tyr Pro Pro Phe Trp Phe Lys Ser Ala Ala Tyr
145 150 155 160

45 Arg Ala Arg Ala Val Arg Asp Pro Arg Gly Leu Ile Ala Glu Trp Gly
165 170 175

50 Thr Met Leu Pro Asp Asp Val Arg Val Arg Val Val Asp Ser Thr Ala
180 185 190

Asp Tyr Arg Trp Met Val Leu Pro Val Arg Pro Ala Gly Thr Ala Gly
195 200 205

55 Trp Asp Glu Glu Arg Leu Ala Ala Ile Val Arg Glu Gly Asp Met Ile
210 215 220

Gly Val Thr Ile Pro Arg Leu
225 230

<211> 576
<212> DNA
<213> Unknown

5 <220>
<223> Metagenome - alpha unit nitrile hydratase - M23dA12

<220>
<221> CDS
10 <222> (1)..(576)
<223>

<400> 47
atg cag ttg cgc gtg cg^g ctg gaa acc gtt cta gcc gaa aag ggt
15 48
Met Gln Leu Arg Val Arg Ala Leu Glu Thr Val Leu Ala Glu Lys Gly
1 5 10 15

20 tat ctc gat ccc gcc gcg ctt gat gcg atg atc gaa gcc tac gag acg
96
Tyr Leu Asp Pro Ala Ala Leu Asp Ala Met Ile Glu Ala Tyr Glu Thr
20 25 30

25 cg^g att ggg ccg cat aac ggc gcg cgc gtc gtc gcc aag gcc tgg tcc
144
Arg Ile Gly Pro His Asn Gly Ala Arg Val Val Ala Lys Ala Trp Ser
35 40 45

30 gac gcc gca ttc aag cga gcg ctg gtc gag gat gcg acc aag gcc gtg
192
Asp Ala Ala Phe Lys Arg Ala Leu Val Glu Asp Ala Thr Lys Ala Val
50 55 60

35 cag tcg ttc ggc gtg gtc aat cgc gtc ggc gat cac ctg atc gcg gtc
240
Gln Ser Phe Gly Val Val Asn Arg Val Gly Asp His Leu Ile Ala Val
65 70 75 80

40 gag aac acg ccc acg ctg cac aac atc atc gtg tgc acg ttg tgc tcc
288
Glu Asn Thr Pro Thr Leu His Asn Ile Ile Val Cys Thr Leu Cys Ser
85 90 95

45 tgc tat ccg tgg gaa gtg ctc ggc ctg ccg ccg gtc tgg tac aaa tcg
336
Cys Tyr Pro Trp Glu Val Leu Gly Leu Pro Pro Val Trp Tyr Lys Ser
100 105 110

50 gcg ccc tac cgc tcg cgc gcg gtc aac gac ccg cgc ggg gta ctc gcc
384
Ala Pro Tyr Arg Ser Arg Ala Val Asn Asp Pro Arg Gly Val Leu Ala
115 120 125

55 gat ttc ggc ctg aag ctg gcg ccg gat atg caa atc cgt gtc tgg gat
432
Asp Phe Gly Leu Lys Leu Ala Pro Asp Met Gln Ile Arg Val Trp Asp
130 135 140

tcg acc gcc gag acg cgc ttc atc gtg ttg ccg atg cgc ccg gcc gga
 480
 Ser Thr Ala Glu Thr Arg Phe Ile Val Leu Pro Met Arg Pro Ala Gly
 145 150 155 160
 5
 acc gac ggc tgg agc gaa gaa aag ctc gcc gcg ctg gtg aca cgc gat
 528
 Thr Asp Gly Trp Ser Glu Glu Lys Leu Ala Ala Leu Val Thr Arg Asp
 165 170 175
 10
 tgc atg atc ggc acc ggc tta ccc aag caa ccc aac gag gtc acg taa
 576
 Cys Met Ile Gly Thr Gly Leu Pro Lys Gln Pro Asn Glu Val Thr
 180 185 190
 15
 <210> 48
 <211> 191
 <212> PRT
 <213> Unknown
 20
 <220>
 <223> Metagenome - alpha unit nitrile hydratase - M23dA12
 25 <400> 48
 Met Gln Leu Arg Val Arg Ala Leu Glu Thr Val Leu Ala Glu Lys Gly
 1 5 10 15
 30
 Tyr Leu Asp Pro Ala Ala Leu Asp Ala Met Ile Glu Ala Tyr Glu Thr
 20 25 30
 35 Arg Ile Gly Pro His Asn Gly Ala Arg Val Val Ala Lys Ala Trp Ser
 35 40 45
 40
 Asp Ala Ala Phe Lys Arg Ala Leu Val Glu Asp Ala Thr Lys Ala Val
 50 55 60
 65 Gln Ser Phe Gly Val Val Asn Arg Val Gly Asp His Leu Ile Ala Val
 70 75 80
 45
 Glu Asn Thr Pro Thr Leu His Asn Ile Ile Val Cys Thr Leu Cys Ser
 85 90 95
 50
 Cys Tyr Pro Trp Glu Val Leu Gly Leu Pro Pro Val Trp Tyr Lys Ser
 100 105 110
 55 Ala Pro Tyr Arg Ser Arg Ala Val Asn Asp Pro Arg Gly Val Leu Ala
 115 120 125
 Asp Phe Gly Leu Lys Leu Ala Pro Asp Met Gln Ile Arg Val Trp Asp

130	135	140
Ser Thr Ala Glu Thr Arg Phe Ile Val Leu Pro Met Arg Pro Ala Gly		
5 145	150	155 160
Thr Asp Gly Trp Ser Glu Glu Lys Leu Ala Ala Leu Val Thr Arg Asp		
165 170 175		
10	Cys Met Ile Gly Thr Gly Leu Pro Lys Gln Pro Asn Glu Val Thr	
	180	185 190
15	<210> 49	
	<211> 624	
	<212> DNA	
	<213> Unknown	
20	<220>	
	<223> Metagenome - alpha unit nitrile hydratase - M49bD9	
25	<220>	
	<221> CDS	
	<222> (1)..(624)	
	<223>	
30	<400> 49	
	atg agc gag cac gat tcc ggc gaa agc cat cac cat ccg cag cca cta	
	48	
	Met Ser Glu His Asp Ser Gly Glu Ser His His His Pro Gln Pro Leu	
	1 5	10 15
35	tcg cag gcg gca ttg cgc gcg aag gcg atc gaa tcg ctg ctg gtc gaa	
	96	
	Ser Gln Ala Ala Leu Arg Ala Lys Ala Ile Glu Ser Leu Leu Val Glu	
	20 25	30
40	aag ggg ctg atc gcg acc gac gtg atc gat cgc gtg gta gca acg tac	
	144	
	Lys Gly Leu Ile Ala Thr Asp Val Ile Asp Arg Val Val Ala Thr Tyr	
	35 40	45
45	gag aaa gaa gtc ggg ccg ctc aac ggc gct aaa gtc gtc gcg cgg gcc	
	192	
	Glu Lys Glu Val Gly Pro Leu Asn Gly Ala Lys Val Val Ala Arg Ala	
	50 55	60
50	tgg acc gat ccg gag tac cgc cgc aga ctg ctc aag aac ggc acg gcg	
	240	
	Trp Thr Asp Pro Glu Tyr Arg Arg Arg Leu Leu Lys Asn Gly Thr Ala	
	65 70	75 80
55	gcg att gcc gag ctg gga ttc ggc ggc ttg cag ggc gaa cac atg atg	
	288	
	Ala Ile Ala Glu Leu Gly Phe Gly Gly Leu Gln Gly Glu His Met Met	
	85 90	95

gtc gtg gaa aac acg ccg tcc gta cat aac gtg atc tgt tgc acg cta
 336
 Val Val Glu Asn Thr Pro Ser Val His Asn Val Ile Cys Cys Thr Leu

100 105

110

5

tgc tca tgc tat ccg tgg ccg gtc ctg gga ctt ccg ccg agc tgg tac
 384
 Cys Ser Cys Tyr Pro Trp Pro Val Leu Gly Leu Pro Pro Ser Trp Tyr
 115 120 125

10

aag tcg ctg gcg tat cgt tcg cga atc gtg cgc gag ccg cgc gcc gtc
 432
 Lys Ser Leu Ala Tyr Arg Ser Arg Ile Val Arg Glu Pro Arg Ala Val
 130 135 140

15

ctc ggc gaa ttc ggc ctc gaa ttg ccc gaa acg gtg gaa gtc cgc gta
 480
 Leu Gly Glu Phe Gly Leu Glu Leu Pro Glu Thr Val Glu Val Arg Val
 145 150 155 160

20

tgg gat agc agt gct gag atg cgc tat ctc gtg ttg ccg gag cgt cca
 528
 Trp Asp Ser Ser Ala Glu Met Arg Tyr Leu Val Leu Pro Glu Arg Pro
 165 170 175

25

gcg gga acg acg gag ttg agc gaa gcg gaa ttg gct tca ttg atc acg
 576
 Ala Gly Thr Thr Glu Leu Ser Glu Ala Glu Leu Ala Ser Leu Ile Thr
 180 185 190

30

cgc gat gcc ttg atc ggc gtg gcg aaa gtc gcg gcg cca agc cgc tag
 624
 Arg Asp Ala Leu Ile Gly Val Ala Lys Val Ala Ala Pro Ser Arg
 195 200 205

35

<210> 50
 <211> 207
 <212> PRT
 <213> Unknown

<220>
 <223> Metagenome - alpha unit nitrile hydratase - M49bD9

45 <400> 50

Met Ser Glu His Asp Ser Gly Glu Ser His His His Pro Gln Pro Leu
 1 5 10 15

50

Ser Gln Ala Ala Leu Arg Ala Lys Ala Ile Glu Ser Leu Leu Val Glu
 20 25 30

55

Lys Gly Leu Ile Ala Thr Asp Val Ile Asp Arg Val Val Ala Thr Tyr
 35 40 45

Glu Lys Glu Val Gly Pro Leu Asn Gly Ala Lys Val Val Ala Arg Ala

040067 AM

30

50

55

60

5 Trp Thr Asp Pro Glu Tyr Arg Arg Arg Leu Leu Lys Asn Gly Thr Ala
65 70 75 80

10 Ala Ile Ala Glu Leu Gly Phe Gly Gly Leu Gln Gly Glu His Met Met
85 90 95

15 val Val Glu Asn Thr Pro Ser Val His Asn Val Ile Cys Cys Thr Leu
100 105 110

20 Cys Ser Cys Tyr Pro Trp Pro Val Leu Gly Leu Pro Pro Ser Trp Tyr
115 120 125

25 Lys Ser Leu Ala Tyr Arg Ser Arg Ile Val Arg Glu Pro Arg Ala Val
130 135 140

30 Leu Gly Glu Phe Gly Leu Glu Leu Pro Glu Thr Val Glu Val Arg Val
145 150 155 160

35 Trp Asp Ser Ser Ala Glu Met Arg Tyr Leu Val Leu Pro Glu Arg Pro
165 170 175

40

Ala Gly Thr Thr Glu Leu Ser Glu Ala Glu Leu Ala Ser Leu Ile Thr
180 185 190

<210> 51
<211> 600
<212> DNA
<213> Unknown

45

<220>
<223> Metagenome - alpha unit nitrile hydratase - M6dE2

50

<220>
<221> CDS
<222> (1)..(600)
<223>

55 <400> 51
atg agc aac cca cgc cgt cga gaa cgg tcg gcc cca ccg gat gcg cga
48 Met Ser Asn Pro Arg Arg Arg Glu Arg Ser Ala Pro Pro Asp Ala Arg
1 5 10 15

gcc aag gcg ctc gca gaa gcg ctt tcg aag caa gga ctc gtg ccg gaa
 96
 Ala Lys Ala Leu Ala Glu Ala Leu Ser Lys Gln Gly Leu Val Pro Glu
 20 25 30
 5
 ggg ttc ctc gac cag gtc ggt tct cac gcc gcg gag gcg tgg agc ccg
 144
 Gly Phe Leu Asp Gln Val Gly Ser His Ala Ala Glu Ala Trp Ser Pro
 35 40 45
 10
 cga aac ggc gca cgg gtc gtg gcg cgg gcg tgg gtg gat ccc gag tac
 192
 Arg Asn Gly Ala Arg Val Val Ala Arg Ala Trp Val Asp Pro Glu Tyr
 50 55 60
 15
 cgg acg cgc ttg ctc gcc gac ggc acc gcc gcg tgc gcc gcg ctc ggc
 240
 Arg Thr Arg Leu Leu Ala Asp Gly Thr Ala Ala Cys Ala Ala Leu Gly
 65 70 75 80
 20
 tac gcg gga ccg cag gga gag tac atc gtg gta ctc gaa gac acg ctg
 288
 Tyr Ala Gly Pro Gln Gly Glu Tyr Ile Val Val Leu Glu Asp Thr Leu
 85 90 95
 25
 gcc gtt cac aac gtg atc gtg tgt acg caa tgc tcg tgt act gcg tgg
 336
 Ala Val His Asn Val Ile Val Cys Thr Gln Cys Ser Cys Thr Ala Trp
 100 105 110
 30
 ccc gtg ctg ggg ctg ccg ccc gat tgg tac aag agt ccg gag tat cgc
 384
 Pro Val Leu Gly Leu Pro Pro Asp Trp Tyr Lys Ser Pro Glu Tyr Arg
 115 120 125
 35
 gcc cgc gtc gtg cgg gag ccg cga cgg gtg ctt cgc gaa atg ggc ctc
 432
 Ala Arg Val Val Arg Glu Pro Arg Arg Val Leu Arg Glu Met Gly Leu
 130 135 140
 40
 gag cta tcc gag agc gtg acg atc cgc gtg tgg gat acg act gcg gaa
 480
 Glu Leu Ser Glu Ser Val Thr Ile Arg Val Trp Asp Thr Thr Ala Glu
 145 150 155 160
 45
 acg cgc ttc ctg gtg ctg ccg ctt cgg ccg gcg gga acc gaa ggg tgg
 528
 Thr Arg Phe Leu Val Leu Pro Leu Arg Pro Ala Gly Thr Glu Gly Trp
 165 170 175
 50
 agc gcg gag cag ctc gcg tcg ctc acg cgc gag gcg atg atc ggc
 576
 Ser Ala Glu Gln Leu Ala Ser Leu Val Thr Arg Glu Ala Met Ile Gly
 180 185 190
 55
 gtg gcg cgg gtc gag gtg gtg tag
 600
 Val Ala Arg Val Glu Val Val
 195

5 <210> 52
 <211> 199
 <212> PRT
 <213> Unknown
 10 <220>
 <223> Metagenome - alpha unit nitrile hydratase - M6dE2
 <400> 52
 Met Ser Asn Pro Arg Arg Arg Glu Arg Ser Ala Pro Pro Asp Ala Arg
 1 5 10 15
 15 Ala Lys Ala Leu Ala Glu Ala Leu Ser Lys Gln Gly Leu Val Pro Glu
 20 25 30
 20 Gly Phe Leu Asp Gln Val Gly Ser His Ala Ala Glu Ala Trp Ser Pro
 35 40 45
 25 Arg Asn Gly Ala Arg Val Val Ala Arg Ala Trp Val Asp Pro Glu Tyr
 50 55 60
 30 Arg Thr Arg Leu Leu Ala Asp Gly Thr Ala Ala Cys Ala Ala Leu Gly
 65 70 75 80
 35 Tyr Ala Gly Pro Gln Gly Glu Tyr Ile Val Val Leu Glu Asp Thr Leu
 85 90 95
 40 Ala Val His Asn Val Ile Val Cys Thr Gln Cys Ser Cys Thr Ala Trp
 100 105 110
 45 Pro Val Leu Gly Leu Pro Pro Asp Trp Tyr Lys Ser Pro Glu Tyr Arg
 115 120 125
 50 Ala Arg Val Val Arg Glu Pro Arg Arg Val Leu Arg Glu Met Gly Leu
 130 135 140
 55 Glu Leu Ser Glu Ser Val Thr Ile Arg Val Trp Asp Thr Thr Ala Glu
 145 150 155 160
 Thr Arg Phe Leu Val Leu Pro Leu Arg Pro Ala Gly Thr Glu Gly Trp
 165 170 175
 55 Ser Ala Glu Gln Leu Ala Ser Leu Val Thr Arg Glu Ala Met Ile Gly
 180 185 190

Val Ala Arg Val Glu Val Val
 195

5
 <210> 53
 <211> 645
 <212> DNA
 <213> Unknown

10
 <220>
 <223> Metagenome - alpha unit nitrile hydratase - M25A18

15
 <220>
 <221> CDS
 <222> (1)..(645)
 <223>

20
 <400> 53
 atg agc ggc acg cat cac cac gac cat gac cac gat cat gac cat gcc
 48
 Met Ser Gly Thr His His Asp His Asp His Asp His Asp His Ala
 1 5 10 15

25
 cat ccg ggc gtc gcc aag gac gag aag gtc cac ggc tat tac caa ttg
 96
 His Pro Gly Val Ala Lys Asp Glu Lys Val His Gly Tyr Tyr Gln Leu
 20 25 30

30
 ctc ggc ctc gcc atc aaa gag ctg ctg atc gaa aaa ggc gtc atc acc
 144
 Leu Gly Leu Ala Ile Lys Glu Leu Leu Ile Glu Lys Gly Val Ile Thr
 35 40 45

35
 gcc gcc gag gtg cgc caa gcg atc gag gcg cgc gac gcg atc acg ccg
 192
 Ala Ala Glu Val Arg Gln Ala Ile Glu Ala Arg Asp Ala Ile Thr Pro
 50 55 60

40
 tcg ctc ggc ggc aag gtg gtc gcg cgc gcc tgg acc gat ccg gcc tac
 240
 Ser Leu Gly Gly Lys Val Val Ala Arg Ala Trp Thr Asp Pro Ala Tyr
 65 70 75 80

45
 aag gcg cgg ctg atc gcc gat ccc gcc gcc atg atg gag atg ggc
 288
 Lys Ala Arg Leu Ile Ala Asp Pro Ala Ala Met Met Glu Met Gly
 85 90 95

50
 gtc gat ctc ggc ccc acc gga ctc gcc atc gcc gag aac acg ccg gag
 336
 Val Asp Leu Gly Pro Thr Gly Leu Ala Ile Ala Glu Asn Thr Pro Glu
 100 105 110

55
 gcg cac aac gtc atc gtc acc ctg tgc tcg tat ccg cgc gcc
 384
 Ala His Asn Val Ile Val Cys Thr Leu Cys Ser Cys Tyr Pro Arg Ala
 115 120 125

gtg ctc ggc ctg ccg ccc tcc tgg tac aag gac cgc gat tac cgg tcg
 432
 Val Leu Gly Leu Pro Pro Ser Trp Tyr Lys Asp Arg Asp Tyr Arg Ser
 130 135 140
 5
 cgc gtg gtg cgc gag ccg cgc gcc gtg ctc aag gag ttc ggc acg gaa
 480
 Arg Val Val Arg Glu Pro Arg Ala Val Leu Lys Glu Phe Gly Thr Glu
 145 150 155 160
 10
 ttg ccc gac gac gtc gac gtc cgc gtc cac gat tcg acc gcc gat ctg
 528
 Leu Pro Asp Asp Val Asp Val Arg Val His Asp Ser Thr Ala Asp Leu
 165 170 175
 15
 cgc tat ctc gtg ctg ccg atg cgc ccg gcc ggc acc gag ggc atg agc
 576
 Arg Tyr Leu Val Leu Pro Met Arg Pro Ala Gly Thr Glu Gly Met Ser
 180 185 190
 20
 gag gcg gag ctg gcc gag atc gtg acg cgc gac tgc atg atc ggc gtg
 624
 Glu Ala Glu Leu Ala Glu Ile Val Thr Arg Asp Cys Met Ile Gly Val
 195 200 205
 25
 acg gtg ccg aaa gcg ccc taa
 645
 Thr Val Pro Lys Ala Pro
 210
 30
 <210> 54
 <211> 214
 <212> PRT
 35 <213> Unknown
 <220>
 <223> Metagenome - alpha unit nitrile hydratase - M25A18
 40 <400> 54
 Met Ser Gly Thr His His Asp His Asp His Asp His Asp His Ala
 1 5 10 15
 45 His Pro Gly Val Ala Lys Asp Glu Lys Val His Gly Tyr Tyr Gln Leu
 20 25 30
 50 Leu Gly Leu Ala Ile Lys Glu Leu Leu Ile Glu Lys Gly Val Ile Thr
 35 40 45
 55 Ala Ala Glu Val Arg Gln Ala Ile Glu Ala Arg Asp Ala Ile Thr Pro
 50 55 60
 Ser Leu Gly Gly Lys Val Val Ala Arg Ala Trp Thr Asp Pro Ala Tyr
 65 70 75 80

Lys Ala Arg Leu Ile Ala Asp Pro Ala Ala Ala Met Met Glu Met Gly
85 90 95

5

Val Asp Leu Gly Pro Thr Gly Leu Ala Ile Ala Glu Asn Thr Pro Glu
100 105 110

10

Ala His Asn Val Ile Val Cys Thr Leu Cys Ser Cys Tyr Pro Arg Ala
115 120 125

15

Val Leu Gly Leu Pro Pro Ser Trp Tyr Lys Asp Arg Asp Tyr Arg Ser
130 135 140

20

Arg Val Val Arg Glu Pro Arg Ala Val Leu Lys Glu Phe Gly Thr Glu
145 150 155 160

25

Leu Pro Asp Asp Val Asp Val Arg Val His Asp Ser Thr Ala Asp Leu
165 170 175

30

Arg Tyr Leu Val Leu Pro Met Arg Pro Ala Gly Thr Glu Gly Met Ser
180 185 190

35

Glu Ala Glu Leu Ala Glu Ile Val Thr Arg Asp Cys Met Ile Gly Val
195 200 205

40 <210> 55

<211> 627

<212> DNA

<213> Unknown

45 <220>

<223> Metagenome - alpha unit nitrile hydratase

50 <220>

<221> CDS

<222> (1)..(627)

<223>

55 <400> 55

atg agc ggt cac cat cac gac cac gac cat gag cac gac aac cac ttc

48

Met Ser Gly His His Asp His Asp His Glu His Asp Asn His Phe

1

5

10

15

96 acg ccg atc gaa gcg cgc gtg aag gcg ctg gaa tcg ctg ctg gtc gcc

Thr Pro Ile Glu Ala Arg Val Lys Ala Leu Glu Ser Leu Leu Val Ala
 20 25 30

aag ggc tat gtc gat ccc gcc gcg ctc gat gcg atc atc gac acc tat
 5 144

Lys Gly Tyr Val Asp Pro Ala Ala Leu Asp Ala Ile Ile Asp Thr Tyr
 35 40 45

gag acg aag atc ggc ccg cgc aac ggc gcc cgc gtc gtc gcc aag gcc
 10 192

Glu Thr Lys Ile Gly Pro Arg Asn Gly Ala Arg Val Val Ala Lys Ala
 50 55 60

tgg acc gat ccg gaa ttc gcg gcg cgg ctc aag cag gat ggc agc gcc
 15 240

Trp Thr Asp Pro Glu Phe Ala Ala Arg Leu Lys Gln Asp Gly Ser Ala
 65 70 75 80

gcc gtc gcc gaa ctc ggc tat ggc ggg cgt ggc ggc gag cat atc gtc
 20 288

Ala Val Ala Glu Leu Gly Tyr Gly Arg Gly Gly Glu His Ile Val
 85 90 95

gcc tgt ttc aat acg ccc gaa gag cac aac ctg atc gtc tgc acg ctc
 25 336

Ala Cys Phe Asn Thr Pro Glu Glu His Asn Leu Ile Val Cys Thr Leu
 100 105 110

tgc tcg tgc tat ccc tgg ccg gtg ctc ggc ctg ccg ccg gtc tgg tac
 30 384

Cys Ser Cys Tyr Pro Trp Pro Val Leu Gly Leu Pro Pro Val Trp Tyr
 115 120 125

aaa tcc ccg ccc tat cgc tcg aaa gcg gtg atc gac ccg cgc ggc gtg
 35 432

Lys Ser Pro Pro Tyr Arg Ser Lys Ala Val Ile Asp Pro Arg Gly Val
 130 135 140

ctg gcc gat ttc ggc gtg acc ctg ccg gag gga caa agg atc cgc gtc
 40 480

Leu Ala Asp Phe Gly Val Thr Leu Pro Glu Gly Gln Arg Ile Arg Val
 145 150 155 160

tgg gat tcc acc gcc gaa acc cgc ttc att gtc atc ccc ctg ccg ccg
 45 528

Trp Asp Ser Thr Ala Glu Thr Arg Phe Ile Val Ile Pro Leu Arg Pro
 165 170 175

gcc ggg acg gaa ggc tgg tcg gaa gaa cag ctg gcg gcg atc gtg acg
 50 576

Ala Gly Thr Glu Gly Trp Ser Glu Glu Gln Leu Ala Ala Ile Val Thr
 180 185 190

cgt gac agc atg atc ggc acc ggc gtg gtc agc gcg gag gct tcg cga
 55 624

Arg Asp Ser Met Ile Gly Thr Gly Val Val Ser Ala Glu Ala Ser Arg
 195 200 205

tga
627

5 <210> 56
<211> 208
<212> PRT
<213> Unknown

10 <220>
<223> Metagenome - alpha unit nitrile hydratase

<400> 56

15 Met Ser Gly His His His Asp His Asp His Glu His Asp Asn His Phe
1 5 10 15

20 Thr Pro Ile Glu Ala Arg Val Lys Ala Leu Glu Ser Leu Leu Val Ala
20 25 30

25 Lys Gly Tyr Val Asp Pro Ala Ala Leu Asp Ala Ile Ile Asp Thr Tyr
35 40 45

30 Glu Thr Lys Ile Gly Pro Arg Asn Gly Ala Arg Val Val Ala Lys Ala
50 55 60

35 Trp Thr Asp Pro Glu Phe Ala Ala Arg Leu Lys Gln Asp Gly Ser Ala
65 70 75 80

40 Ala Val Ala Glu Leu Gly Tyr Gly Gly Arg Gly Gly Glu His Ile Val
85 90 95

45 Ala Cys Phe Asn Thr Pro Glu Glu His Asn Leu Ile Val Cys Thr Leu
100 105 110

50 Cys Ser Cys Tyr Pro Trp Pro Val Leu Gly Leu Pro Pro Val Trp Tyr
115 120 125

55 Lys Ser Pro Pro Tyr Arg Ser Lys Ala Val Ile Asp Pro Arg Gly Val
130 135 140

60 Leu Ala Asp Phe Gly Val Thr Leu Pro Glu Gly Gln Arg Ile Arg Val
145 150 155 160

65 Trp Asp Ser Thr Ala Glu Thr Arg Phe Ile Val Ile Pro Leu Arg Pro
165 170 175

70 Ala Gly Thr Glu Gly Trp Ser Glu Glu Gln Leu Ala Ala Ile Val Thr

180

185

190

5 Arg Asp Ser Met Ile Gly Thr Gly Val Val Ser Ala Glu Ala Ser Arg
 195 200 205

10 <210> 57

<211> 696

<212> DNA

<213> Unknown

15 <220>

<223> Metagenome - alpha unit nitrile hydratase - M3aG10

20 <220>

<221> CDS

<222> (1)..(696)

<223>

25 <400> 57

atg gat cca acg agg cgt agt ttc ctg gcg tct acc gtt gcc ctg acc
 48

Met Asp Pro Thr Arg Arg Ser Phe Leu Ala Ser Thr Val Ala Leu Thr
 1 5 10 15

ggc ggc gca gct atc ccc gat ctg gct cat gcg gca gac cac gac cac
 96

Gly Gly Ala Ala Ile Pro Asp Leu Ala His Ala Ala Asp His Asp His
 20 25 30

cag cat caa gat ttg ccg tcc gat ctg gcg ctg cgg gtg aag tcg ctc
 144

Gln His Gln Asp Leu Pro Ser Asp Leu Ala Leu Arg Val Lys Ser Leu
 35 35 40 45

gaa tcg ctg ctt gtc gag aag ggg ctg gtg gag cga gca gcg ctc gac
 192

Glu Ser Leu Leu Val Glu Lys Gly Leu Val Glu Arg Ala Ala Leu Asp
 50 55 60

gcg ctg gtg gac acc tac gag cac aaa gtc ggg ccg cga aac gga gcg
 240

Ala Leu Val Asp Thr Tyr Glu His Lys Val Gly Pro Arg Asn Gly Ala
 65 70 75 80

cgc gtt gtc gcg ccg gcc tgg gtt gac ccg gac tac aag caa ccg tta
 288

Arg Val Val Ala Arg Ala Trp Val Asp Pro Asp Tyr Lys Gln Arg Leu
 85 90 95

ttc gcg aac ggt acc gcc gca gtc gcg gag ttc ggc tac tcc ggc tcg
 336

Phe Ala Asn Gly Thr Ala Ala Val Ala Glu Phe Gly Tyr Ser Gly Ser
 100 105 110

cag ggc gct gac atc ccg gtc gaa aac acg gcc act gta cat aac
 384

Gln Gly Ala Asp Ile Arg Val Val Glu Asn Thr Ala Thr Val His Asn

	115	120	125
	ctc gtc gtg tgc acg ctg tgc tct tgt tat ccc tgg ccg gtg ctg ggc		
432			
5	Leu Val Val Cys Thr Leu Cys Ser Cys Tyr Pro Trp Pro Val Leu Gly		
	130	135	140
	ttg ccg ccg gtc tgg tac aag tcc gcg ccc tat cgg tct cgc gtg gtg		
480			
10	Leu Pro Pro Val Trp Tyr Lys Ser Ala Pro Tyr Arg Ser Arg Val Val		
	145	150	155
	atc gat ccg cga ggt gtg ctg cgc gag ttc ggc gtg gtg ctg ccg gac		
528			
15	Ile Asp Pro Arg Gly Val Leu Arg Glu Phe Gly Val Val Leu Pro Asp		
	165	170	175
	cat atc gaa gtg cgt gtc tat gac agc acg gcg gag caa cgc tat cta		
576			
20	His Ile Glu Val Arg Val Tyr Asp Ser Thr Ala Glu Gln Arg Tyr Leu		
	180	185	190
	gtg ctg ccg gag cgg ccg gcc gga acc gaa aac ctg aca gaa gcg		
624			
25	Val Leu Pro Glu Arg Pro Ala Gly Thr Glu Asn Leu Thr Glu Glu Ala		
	195	200	205
	ctg gcg ctg ctg gtg acg cgc gac gcg atg att ggc gtg gcc aag gtc		
672			
30	Leu Ala Leu Leu Val Thr Arg Asp Ala Met Ile Gly Val Ala Lys Val		
	210	215	220
	gcg ccg ccg gga ggc cgc gga tga		
696			
35	Ala Pro Pro Gly Gly Arg Gly		
	225	230	
	<210> 58		
40	<211> 231		
	<212> PRT		
	<213> Unknown		
	<220>		
45	<223> Metagenome - alpha unit nitrile hydratase - M3aG10		
	<400> 58		
	Met Asp Pro Thr Arg Arg Ser Phe Leu Ala Ser Thr Val Ala Leu Thr		
50	1	5	10
	15		
	Gly Gly Ala Ala Ile Pro Asp Leu Ala His Ala Ala Asp His Asp His		
	20	25	30
55	Gln His Gln Asp Leu Pro Ser Asp Leu Ala Leu Arg Val Lys Ser Leu		
	35	40	45

Glu Ser Leu Leu Val Glu Lys Gly Leu Val Glu Arg Ala Ala Leu Asp
50 55 60

5 Ala Leu Val Asp Thr Tyr Glu His Lys Val Gly Pro Arg Asn Gly Ala
65 70 75 80

10 Arg Val Val Ala Arg Ala Trp Val Asp Pro Asp Tyr Lys Gln Arg Leu
85 90 95

15 Phe Ala Asn Gly Thr Ala Ala Val Ala Glu Phe Gly Tyr Ser Gly Ser
100 105 110

20 Gln Gly Ala Asp Ile Arg Val Val Glu Asn Thr Ala Thr Val His Asn
115 120 125

25 Leu Val Val Cys Thr Leu Cys Ser Cys Tyr Pro Trp Pro Val Leu Gly
130 135 140

30 His Ile Glu Val Arg Val Leu Arg Glu Phe Gly Val Val Leu Pro Asp
165 170 175

35 Val Leu Pro Glu Arg Pro Ala Gly Thr Glu Asn Leu Thr Glu Glu Ala
195 200 205

40 Leu Ala Leu Leu Val Thr Arg Asp Ala Met Ile Gly Val Ala Lys Val
210 215 220

45 Ala Pro Pro Gly Gly Arg Gly
225 230

50 <210> 59
<211> 609
<212> DNA
<213> Unknown

55 <220>
<223> Metagenome - alpha unit nitrile hydratase - M73dC9
<220>
<221> CDS

<222> (1)...(609)
 <223>
 <400> 59
 5 atg agc tcg aag ccc acc gaa gat ctc ggc acc tac cag ccg ctc acc
 48 Met Ser Ser Lys Pro Thr Glu Asp Leu Gly Thr Tyr Gln Pro Leu Thr
 1 5 10 15
 10 tac tac cag atg atg gaa gtg agc ctg cgc gag ctg ctg gtg gag aag
 96 Tyr Tyr Gln Met Met Glu Val Ser Leu Arg Glu Leu Leu Val Glu Lys
 20 25 30
 15 ggc gtg atc acc gaa gcg gaa gtc gcc cgc gcg atg ggc gag atc ggc
 144 Gly Val Ile Thr Glu Ala Glu Val Ala Arg Ala Met Gly Glu Ile Gly
 35 40 45
 20 gcg aga agc ccg gag cgc ggc gcg aag atg gtc gcg cgc gcg tgg gtg
 192 Ala Arg Ser Pro Glu Arg Gly Ala Lys Met Val Ala Arg Ala Trp Val
 50 55 60
 25 gac ccg gcg tac aag gcg cgc atg ctt gcc gac ggc agc aag gcc gcc
 240 Asp Pro Ala Tyr Lys Ala Arg Met Leu Ala Asp Gly Ser Lys Ala Ala
 65 70 75 80
 30 gag gag ctc ggg ttc gag gtg ccg ggc ctc aag ctg atc gtg gtc gag
 288 Glu Glu Leu Gly Phe Glu Val Pro Gly Leu Lys Leu Ile Val Val Glu
 85 90 95
 35 aac acc gcg gac acg cac aac gtg gtc gtg tgc acg ctg tgc tcg tgc
 336 Asn Thr Ala Asp Thr His Asn Val Val Val Cys Thr Leu Cys Ser Cys
 100 105 110
 40 tac ccg cgc atc ctg ctc ggc atc ccg ccc gag tgg tac aag tcg cgc
 384 Tyr Pro Arg Ile Leu Leu Gly Ile Pro Pro Glu Trp Tyr Lys Ser Arg
 115 120 125
 45 agc tac cgc agc cgc aca gtg cgc gag ccg cgc gcg gtg ctc gcc gaa
 432 Ser Tyr Arg Ser Arg Thr Val Arg Glu Pro Arg Ala Val Leu Ala Glu
 130 135 140
 50 ttc ggc acg acc atc ccg gag aac gtc gcg atc cga gtg cac gac agc
 480 Phe Gly Thr Thr Ile Pro Glu Asn Val Ala Ile Arg Val His Asp Ser
 145 150 155 160
 55 act gcg gac atg cgc tac ctc gtg atg ccg atg ccg cct gcg ggc acc
 528 Thr Ala Asp Met Arg Tyr Leu Val Met Pro Met Arg Pro Ala Gly Thr
 165 170 175

gaa aat ttc acc gaa gag cag ctc gct gca ttg gtg acg cgc gac agc
576
Glu Asn Phe Thr Glu Glu Gln Leu Ala Ala Leu Val Thr Arg Asp Ser
180 185 190

5 ctg atc ggt gtt tcc tta gca acg ctt ccg tag
609
Leu Ile Gly Val Ser Leu Ala Thr Leu Pro
195 200

10

<210> 60
<211> 202
<212> PRT
15 <213> Unknown

<220>
<223> Metagenome - alpha unit nitrile hydratase - M73dC9

20 <400> 60

Met Ser Ser Lys Pro Thr Glu Asp Leu Gly Thr Tyr Gln Pro Leu Thr
1 5 10 15

25 Tyr Tyr Gln Met Met Glu Val Ser Leu Arg Glu Leu Leu Val Glu Lys
20 25 30

30 Gly Val Ile Thr Glu Ala Glu Val Ala Arg Ala Met Gly Glu Ile Gly
35 40 45

35 Ala Arg Ser Pro Glu Arg Gly Ala Lys Met Val Ala Arg Ala Trp Val
50 55 60

40 Asp Pro Ala Tyr Lys Ala Arg Met Leu Ala Asp Gly Ser Lys Ala Ala
65 70 75 80

45 Glu Glu Leu Gly Phe Glu Val Pro Gly Leu Lys Leu Ile Val Val Glu
85 90 95

50 Tyr Pro Arg Ile Leu Leu Gly Ile Pro Pro Glu Trp Tyr Lys Ser Arg
115 120 125

55 Ser Tyr Arg Ser Arg Thr Val Arg Glu Pro Arg Ala Val Leu Ala Glu
130 135 140

Phe Gly Thr Thr Ile Pro Glu Asn Val Ala Ile Arg Val His Asp Ser
145 150 155 160

Thr Ala Asp Met Arg Tyr Leu Val Met Pro Met Arg Pro Ala Gly Thr
165 170 175

5

Glu Asn Phe Thr Glu Glu Gln Leu Ala Ala Leu Val Thr Arg Asp Ser
180 185 190

10

Leu Ile Gly Val Ser Leu Ala Thr Leu Pro
195 200

15 <210> 61
<211> 825
<212> DNA
<213> Unknown

20 <220>
<223> Metagenome - beta unit nitrile hydratase

<220>
<221> CDS
25 <222> (1) .. (825)
<223>

<400> 61
atg gtg gga cgt ggg aag tgg gca ctt ggc agt agg cag ttt gct gcg
30 48
Met Val Gly Arg Gly Lys Trp Ala Leu Gly Ser Arg Gln Phe Ala Ala
1 5 10 15

35 gct gcc aac tgg caa ctt atc agt cgc cct tca tgg tca gct tgt aat
96
Ala Ala Asn Trp Gln Leu Ile Ser Arg Pro Ser Trp Ser Ala Cys Asn
20 25 30

40 ata ttg gtc ctc atg agc gcc acg cac ccc aaa aag cgc gcc gac
144
Ile Leu Val Leu Met Ser Ala Thr His Pro Lys Lys Arg Ala Ala Asp
35 40 45

45 atc ggc ggc aac aaa gcc ggc gcg gtg gac acc gcg gat cac ggc atg
192
Ile Gly Gly Asn Lys Ala Gly Ala Val Asp Thr Ala Asp His Gly Met
50 55 60

50 aag ttc tgg gag cgg cag gcc aac gcc ctg cgc acc gcg ctg cgg cgc
240
Lys Phe Trp Glu Arg Gln Ala Asn Ala Leu Arg Thr Ala Leu Arg Arg
65 70 75 80

55 aat gga ctg atg agc gta gat gag ctg cgc cgc gca gcg gag gac ctg
288
Asn Gly Leu Met Ser Val Asp Glu Leu Arg Arg Ala Ala Glu Asp Leu
85 90 95

gga gac cgc tac gcg aag ctt gag tac ttc gag cgc acg acg ttc gcg
 336
 Gly Asp Arg Tyr Ala Lys Leu Glu Tyr Phe Glu Arg Thr Thr Phe Ala
 100 105 110
 5
 ctg cgc acg gtc ctg ctc gaa aag ggc tac ttc acg gag gag tcg ctc
 384
 Leu Arg Thr Val Leu Leu Glu Lys Gly Tyr Phe Thr Glu Glu Ser Leu
 115 120 125
 10
 gcg gcg aag atg gcc gag gtg cgg aag ccg ctt cga tgt gcc gcg caa
 432
 Ala Ala Lys Met Ala Glu Val Arg Lys Pro Leu Arg Cys Ala Ala Gln
 130 135 140
 15
 gaa gga att gcc ggt gaa gaa agt gaa gcg atg aac cca gcg acg
 480
 Glu Gly Ile Ala Gly Glu Glu Ser Glu Ala Met Asn Pro Ala Thr
 145 150 155 160
 20
 ggc aag cag gac ggc caa cgg ctg cca tct acg tat acc gcg gcg ccc
 528
 Gly Lys Gln Asp Gly Gln Arg Leu Pro Ser Thr Tyr Thr Ala Ala Pro
 165 170 175
 25
 ggg cac cga ttc gat gtc ggt gac cgc gtt gtg gtc aag cgc tca aat
 576
 Gly His Arg Phe Asp Val Gly Asp Arg Val Val Val Lys Arg Ser Asn
 180 185 190
 30
 ccg ccc ggc cac cgc cgc acg cct cat tac atc cgc ggc aag acg ggc
 624
 Pro Pro Gly His Arg Arg Thr Pro His Tyr Ile Arg Gly Lys Thr Gly
 195 200 205
 35
 gtg atc gag cgc atc tgc ggc gcc ttc ccc aac ccg gaa gag ctg gca
 672
 Val Ile Glu Arg Ile Cys Gly Ala Phe Pro Asn Pro Glu Glu Leu Ala
 210 215 220
 40
 tac gga ttc gac ggc gaa ccg aag aag gtg ctc tac cgc gtg cga ttc
 720
 Tyr Gly Phe Asp Gly Glu Pro Lys Lys Val Leu Tyr Arg Val Arg Phe
 225 230 235 240
 45
 cgg caa aaa gag gtg tgg ccg tat cgc ggc ccg gcg cac gac gtg
 768
 Arg Gln Lys Glu Val Trp Pro Ala Tyr Arg Gly Pro Ala His Asp Val
 245 250 255
 50
 atc gag atg gag att ttc gag cat tgg ctc gag ccg gcg cag agc cag
 816
 Ile Glu Met Glu Ile Phe Glu His Trp Leu Glu Pro Ala Gln Ser Gln
 260 265 270
 55
 aaa acc tga
 825
 Lys Thr

5 <210> 62
<211> 274
<212> PRT
<213> Unknown

10 <220>
<223> Metagenome - beta unit nitrile hydratase

15 <400> 62

Met Val Gly Arg Gly Lys Trp Ala Leu Gly Ser Arg Gln Phe Ala Ala
1 5 10 15

20 Ala Ala Asn Trp Gln Leu Ile Ser Arg Pro Ser Trp Ser Ala Cys Asn
20 25 30

25 Ile Leu Val Leu Met Ser Ala Thr His Pro Lys Lys Arg Ala Ala Asp
35 40 45

30 Ile Gly Gly Asn Lys Ala Gly Ala Val Asp Thr Ala Asp His Gly Met
50 55 60

35 Lys Phe Trp Glu Arg Gln Ala Asn Ala Leu Arg Thr Ala Leu Arg Arg
65 70 75 80

40 Asn Gly Leu Met Ser Val Asp Glu Leu Arg Arg Ala Ala Glu Asp Leu
85 90 95

45 Gly Asp Arg Tyr Ala Lys Leu Glu Tyr Phe Glu Arg Thr Thr Phe Ala
100 105 110

50 Leu Arg Thr Val Leu Leu Glu Lys Gly Tyr Phe Thr Glu Glu Ser Leu
115 120 125

55 Ala Ala Lys Met Ala Glu Val Arg Lys Pro Leu Arg Cys Ala Ala Gln
130 135 140

60 Glu Gly Ile Ala Gly Glu Glu Glu Ser Glu Ala Met Asn Pro Ala Thr
145 150 155 160

65 Gly Lys Gln Asp Gly Gln Arg Leu Pro Ser Thr Tyr Thr Ala Ala Pro
165 170 175

70 Gly His Arg Phe Asp Val Gly Asp Arg Val Val Val Lys Arg Ser Asn
180 185 190

Pro Pro Gly His Arg Arg Thr Pro His Tyr Ile Arg Gly Lys Thr Gly
 195 200 205

5 Val Ile Glu Arg Ile Cys Gly Ala Phe Pro Asn Pro Glu Glu Leu Ala
 210 215 220

10 Tyr Gly Phe Asp Gly Glu Pro Lys Lys Val Leu Tyr Arg Val Arg Phe
 225 230 235 240

15 Arg Gln Lys Glu Val Trp Pro Ala Tyr Arg Gly Pro Ala His Asp Val
 245 250 255

20 Ile Glu Met Glu Ile Phe Glu His Trp Leu Glu Pro Ala Gln Ser Gln
 260 265 270

Lys Thr

25

<210> 63
 <211> 627
 <212> DNA
 <213> Unknown

30

<220>
 <223> Metagenome - beta unit nitrile hydratase - M12K24

35

<220>
 <221> CDS
 <222> (1)..(627)
 <223>

40

<400> 63
 atg gac ggc atg cac gac ctg gga ggc agg cag ggc ttc gga ccg gtt
 48
 Met Asp Gly Met His Asp Leu Gly Gly Arg Gln Gly Phe Gly Pro Val
 1 5 10 15

45

cgc tac acg atc gac gcg ccc gca ttc cat tcg ccg tgg gaa gtg cgc
 96
 Arg Tyr Thr Ile Asp Ala Pro Ala Phe His Ser Pro Trp Glu Val Arg
 20 25 30

50

gcg aat tcg ctc tat gcg ttc gcg gtg cgc ctc ggc atc ttc aac atg
 144
 Ala Asn Ser Leu Tyr Ala Phe Ala Val Arg Leu Gly Ile Phe Asn Met
 35 40 45

55

gac gaa tac cgc cat gcg atc gag cgg atg gag ccg cgc cat tac ctc
 192
 Asp Glu Tyr Arg His Ala Ile Glu Arg Met Glu Pro Arg His Tyr Leu
 50 55 60

ggc gcc ggc tat tac gaa cgc tcg ttg acc ggc ctc gcg acc ttg ctg
 240
 Gly Ala Gly Tyr Tyr Glu Arg Ser Leu Thr Gly Leu Ala Thr Leu Leu
 65 70 75 80
 5
 gtc gag aag ggc gtc gtg acg cgc gag gaa ctc gag acc cgg gcg cag
 288
 Val Glu Lys Gly Val Val Thr Arg Glu Glu Leu Glu Thr Arg Ala Gln
 85 90 95
 10
 ggc cgc tac ccg ctg gcg atg ccc agc gcg cct ggc cgc acc aat gcg
 336
 Gly Arg Tyr Pro Leu Ala Met Pro Ser Ala Pro Gly Arg Thr Asn Ala
 100 105 110
 15
 cag gca cgc gag cgt ttc cag ccg ggc gac cgg gtt cgc gtc aag gcg
 384
 Gln Ala Arg Glu Arg Phe Gln Pro Gly Asp Arg Val Arg Val Lys Ala
 115 120 125
 20
 gat ttc gtg tcg ggg cac gtg cgg atg ccg gcg tac atc cgc ggc aag
 432
 Asp Phe Val Ser Gly His Val Arg Met Pro Ala Tyr Ile Arg Gly Lys
 130 135 140
 25
 acc ggc gtg gtc agc gag tcc ccg gac tat ccg ttt ccc gat gcg
 480
 Thr Gly Val Val Val Ser Glu Ser Pro Asp Tyr Pro Phe Pro Asp Ala
 145 150 155 160
 30
 cat gcg cac tcg gtc gat gcc cag gac gag cca acc tac gac gtg cgc
 528
 His Ala His Ser Val Asp Ala Gln Asp Glu Pro Thr Tyr Asp Val Arg
 165 170 175
 35
 ttc cgc agc gag gat cta tgg ccg gat tcc gcc gat tcc gca ctc gtt
 576
 Phe Arg Ser Glu Asp Leu Trp Pro Asp Ser Ala Asp Ser Ala Leu Val
 180 185 190
 40
 cac gtc ggc gta ttc cag agc tac ctc gag cgg gag tcg acg cca gga
 624
 His Val Gly Val Phe Gln Ser Tyr Leu Glu Arg Glu Ser Thr Pro Gly
 195 200 205
 45
 tag
 627
 50 <210> 64
 <211> 208
 <212> PRT
 <213> Unknown
 55 <220>
 <223> Metagenome - beta unit nitrile hydratase - M12K24
 <400> 64

Met Asp Gly Met His Asp Leu Gly Gly Arg Gln Gly Phe Gly Pro Val
1 5 10 15

5 Arg Tyr Thr Ile Asp Ala Pro Ala Phe His Ser Pro Trp Glu Val Arg
20 25 30

10 Ala Asn Ser Leu Tyr Ala Phe Ala Val Arg Leu Gly Ile Phe Asn Met
35 40 45

15 Asp Glu Tyr Arg His Ala Ile Glu Arg Met Glu Pro Arg His Tyr Leu
50 55 60

20 Gly Ala Gly Tyr Tyr Glu Arg Ser Leu Thr Gly Leu Ala Thr Leu Leu
65 70 75 80

25 Val Glu Lys Gly Val Val Thr Arg Glu Glu Leu Glu Thr Arg Ala Gln
85 90 95

30 Gly Arg Tyr Pro Leu Ala Met Pro Ser Ala Pro Gly Arg Thr Asn Ala
100 105 110

35 Gln Ala Arg Glu Arg Phe Gln Pro Gly Asp Arg Val Arg Val Lys Ala
115 120 125

40 Asp Phe Val Ser Gly His Val Arg Met Pro Ala Tyr Ile Arg Gly Lys
130 135 140

45 Thr Gly Val Val Val Ser Glu Ser Pro Asp Tyr Pro Phe Pro Asp Ala
145 150 155 160

50 His Ala His Ser Val Asp Ala Gln Asp Glu Pro Thr Tyr Asp Val Arg
165 170 175

55 Phe Arg Ser Glu Asp Leu Trp Pro Asp Ser Ala Asp Ser Ala Leu Val
180 185 190

60 His Val Gly Val Phe Gln Ser Tyr Leu Glu Arg Glu Ser Thr Pro Gly
195 200 205

<210> 65

<211> 660

<212> DNA

<213> Unknown

<220>

<223> Metagenome - beta unit nitrile hydratase - M29M24

<220>
 <221> CDS
 <222> (1)..(660)
 5 <223>

 <400> 65
 atg aac ggc gtt cat gac atg ggc ggc atg cac ggc atg ggt gcg atc
 48 Met Asn Gly Val His Asp Met Gly Gly Met His Gly Met Gly Ala Ile
 10 1 5 10 15

 cgc cgc gag gag aac gag ccc gct ttc cac gag ccc tgg gag ggg cgg
 96 Arg Arg Glu Glu Asn Glu Pro Ala Phe His Glu Pro Trp Glu Gly Arg
 15 20 25 30

 gtt ttc gct ctg acc acg gcg gtc gag gcc tgg ggt cgg tgg acc ctc
 144 Val Phe Ala Leu Thr Thr Ala Val Glu Ala Trp Gly Arg Trp Thr Leu
 20 35 40 45

 gat gct tcc cga cac cgg atc gag cgg atg aat gcg gcg gac tat ctg
 192 Asp Ala Ser Arg His Arg Ile Glu Arg Met Asn Ala Ala Asp Tyr Leu
 25 50 55 60

 cgg gtg agc tat tac gag aag tgg ctc gag tcg ctt ctc gct ctc ctg
 240 Arg Val Ser Tyr Tyr Glu Lys Trp Leu Glu Ser Leu Leu Ala Leu Leu
 30 65 70 75 80

 tcc gag acc gga atg gcg agt ccg gcg gag ata cgg agt ggg gag cgt
 288 Ser Glu Thr Gly Met Ala Ser Pro Ala Glu Ile Arg Ser Gly Glu Arg
 35 85 90 95

 gcc gac ggc aca ccg aaa gcg acc ccg ccg ctc ccg gcc gac cac gtg
 336 Ala Asp Gly Thr Pro Lys Ala Thr Pro Pro Leu Pro Ala Asp His Val
 40 100 105 110

 acg gcg att ctc gcc agc ggg ttt ccc gcg agc cgg gag gcg gga gct
 384 Thr Ala Ile Leu Ala Ser Gly Phe Pro Ala Ser Arg Glu Ala Gly Ala
 45 115 120 125

 gcg cct cgc ttc cga gtg agc gag cgg gtg cgc acc cgg aac atc aac
 432 Ala Pro Arg Phe Arg Val Ser Glu Arg Val Arg Thr Arg Asn Ile Asn
 50 130 135 140

 ccg acg act cac acg cgc ctt ccg cga tac gcc cgg cgg aag ctc ggg
 480 Pro Thr Thr His Thr Arg Leu Pro Arg Tyr Ala Arg Arg Lys Leu Gly
 55 145 150 155 160

 acg atc gag cgc gac cac gga gtg ttc gtc ttc ccg gat acg aac gcg
 528

Thr Ile Glu Arg Asp His Gly Val Phe Val Phe Pro Asp Thr Asn Ala
 165 170 175
 cac gct ctc ggg gag aaa ccg cag cac gtc tat tcg gtt cgt ttc gag
 5 576
 His Ala Leu Gly Glu Lys Pro Gln His Val Tyr Ser Val Arg Phe Glu
 180 185 190
 gcg cgt gag ctc tgg ggc gag act gcc agg cca gag gat tcc gtc tac
 10 624
 Ala Arg Glu Leu Trp Gly Glu Thr Ala Arg Pro Glu Asp Ser Val Tyr
 195 200 205
 atc gat ctt tgg gac gag tac ctt gaa ccc gtg tag
 15 660
 Ile Asp Leu Trp Asp Glu Tyr Leu Glu Pro Val
 210 215

20 <210> 66
 <211> 219
 <212> PRT
 <213> Unknown

25 <220>
 <223> Metagenome - beta unit nitrile hydratase - M29M24

<400> 66

30 Met Asn Gly Val His Asp Met Gly Gly Met His Gly Met Gly Ala Ile
 1 5 10 15

Arg Arg Glu Glu Asn Glu Pro Ala Phe His Glu Pro Trp Glu Gly Arg
 35 20 25 30

Val Phe Ala Leu Thr Thr Ala Val Glu Ala Trp Gly Arg Trp Thr Leu
 40 35 40 45

40 Asp Ala Ser Arg His Arg Ile Glu Arg Met Asn Ala Ala Asp Tyr Leu
 50 55 60

45 Arg Val Ser Tyr Tyr Glu Lys Trp Leu Glu Ser Leu Leu Ala Leu Leu
 65 70 75 80

50 Ser Glu Thr Gly Met Ala Ser Pro Ala Glu Ile Arg Ser Gly Glu Arg
 85 90 95

55 Ala Asp Gly Thr Pro Lys Ala Thr Pro Pro Leu Pro Ala Asp His Val
 100 105 110

Thr Ala Ile Leu Ala Ser Gly Phe Pro Ala Ser Arg Glu Ala Gly Ala
 115 120 125

Ala Pro Arg Phe Arg Val Ser Glu Arg Val Arg Thr Arg Asn Ile Asn
130 135 140

5 Pro Thr Thr His Thr Arg Leu Pro Arg Tyr Ala Arg Arg Lys Leu Gly
145 150 155 160

10 Thr Ile Glu Arg Asp His Gly Val Phe Val Phe Pro Asp Thr Asn Ala
165 170 175

15 His Ala Leu Gly Glu Lys Pro Gln His Val Tyr Ser Val Arg Phe Glu
180 185 190

20 Ala Arg Glu Leu Trp Gly Glu Thr Ala Arg Pro Glu Asp Ser Val Tyr
195 200 205

25 Ile Asp Leu Trp Asp Glu Tyr Leu Glu Pro Val
210 215

30 <210> 67
<211> 660
<212> DNA
<213> Unknown

35 <220>
<221> CDS
<222> (1)..(660)
<223>

40 <400> 67
atg acc aat tcg ctg cac gac atg ggc ggc atg cac ggc ttt ggc cgg
48 Met Thr Asn Ser Leu His Asp Met Gly Gly Met His Gly Phe Gly Arg
1 5 10 15

45 gtc gag ccc gag ccg aac gag ccg ccg ttt cac cag cgc tgg gag ggc
96 Val Glu Pro Glu Pro Asn Glu Pro Pro Phe His Gln Arg Trp Glu Gly
20 25 30

50 cgg gtg ctg ggg atg cag cgc gcc atg ggc ttt acc ggg ctg tgg acc
144 Arg Val Leu Gly Met Gln Arg Ala Met Gly Phe Thr Gly Leu Trp Thr
35 40 45

55 atc gac gcc ggc cgc gcc tcg ctc gaa gcc ctg ccg cca tta gcg tat
192 Ile Asp Ala Gly Arg Ala Ser Leu Glu Ala Leu Pro Pro Leu Ala Tyr
50 55 60

ctg ggt tcg tcc tac tat cgg cgc tgg ttt ctt ggc ctg gag agc cgg
 240
 5 Leu Gly Ser Ser Tyr Tyr Arg Arg Trp Phe Leu Gly Leu Glu Ser Arg
 65 70 75 80
 ctg ctg ctg cgc ggc ctc gtt ggc gag gac gag atc gcg gca ggc cgt
 288
 10 Leu Leu Leu Arg Gly Leu Val Gly Glu Asp Glu Ile Ala Ala Gly Arg
 85 90 95
 tcg atg cgc gcc ggc gcc atg ttg ccg cgc acc ctg acc cag gcc gat
 336
 15 Ser Met Arg Ala Gly Ala Met Leu Pro Arg Thr Leu Thr Gln Ala Asp
 100 105 110
 gtg gag aaa acc ctg acc cgc ggc gac ttc gcc cgc ccg acc aac acc
 384
 20 Val Glu Lys Thr Leu Thr Arg Gly Asp Phe Ala Arg Pro Thr Asn Thr
 115 120 125
 ccg gcg cgt ttc cag ccg ggc gac cgg gtg caa acg aag aac atc aac
 432
 25 Pro Ala Arg Phe Gln Pro Gly Asp Arg Val Gln Thr Lys Asn Ile Asn
 130 135 140
 ccg gcg acc cac acc cgc ctg ccg cgc tat gcc cgc ggc aag act ggc
 480
 30 Pro Ala Thr His Thr Arg Leu Pro Arg Tyr Ala Arg Gly Lys Thr Gly
 145 150 155 160
 acg gtc gag gcg gtc cgc ggc gtt cac gtc ttt ccc gac acc gcc gcg
 528
 35 Thr Val Glu Ala Val Arg Gly Val His Val Phe Pro Asp Thr Ala Ala
 165 170 175
 ctc ggc gcc ggc gac gac ccg caa tgg ctc tac gcc gtg gtc ttc ccg
 576
 40 Leu Gly Ala Gly Asp Asp Pro Gln Trp Leu Tyr Ala Val Val Phe Pro
 180 185 190
 gcg cgc gag ttg tgg gga gag gcg gcc gat ccc gcg atc aaa atc tgc
 624
 45 Ala Arg Glu Leu Trp Gly Glu Ala Ala Asp Pro Ala Ile Lys Ile Ser
 195 200 205
 atc gag gcg ttc gaa ccc tat atc gac ccc gca tga
 660
 50 Ile Glu Ala Phe Glu Pro Tyr Ile Asp Pro Ala
 210 215
 <210> 68
 <211> 219
 55 <212> PRT
 <213> Unknown
 <220>
 <223> Metagenome - beta unit nitrile hydratase - M2K17

<400> 68

5 Met Thr Asn Ser Leu His Asp Met Gly Gly Met His Gly Phe Gly Arg
1 5 10 15

10 Val Glu Pro Glu Pro Asn Glu Pro Pro Phe His Gln Arg Trp Glu Gly
20 25 30

15 Arg Val Leu Gly Met Gln Arg Ala Met Gly Phe Thr Gly Leu Trp Thr
35 40 45

20 Ile Asp Ala Gly Arg Ala Ser Leu Glu Ala Leu Pro Pro Leu Ala Tyr
50 55 60

25 Leu Gly Ser Ser Tyr Tyr Arg Arg Trp Phe Leu Gly Leu Glu Ser Arg
65 70 75 80

30 Leu Leu Leu Arg Gly Leu Val Gly Glu Asp Glu Ile Ala Ala Gly Arg
85 90 95

35 Ser Met Arg Ala Gly Ala Met Leu Pro Arg Thr Leu Thr Gln Ala Asp
100 105 110

40 Val Glu Lys Thr Leu Thr Arg Gly Asp Phe Ala Arg Pro Thr Asn Thr
115 120 125

45 Pro Ala Arg Phe Gln Pro Gly Asp Arg Val Gln Thr Lys Asn Ile Asn
130 135 140

50 Pro Ala Thr His Thr Arg Leu Pro Arg Tyr Ala Arg Gly Lys Thr Gly
145 150 155 160

55 Thr Val Glu Ala Val Arg Gly Val His Val Phe Pro Asp Thr Ala Ala
165 170 175

Leu Gly Ala Gly Asp Asp Pro Gln Trp Leu Tyr Ala Val Val Phe Pro
180 185 190

55 Ala Arg Glu Leu Trp Gly Glu Ala Ala Asp Pro Ala Ile Lys Ile Ser
195 200 205

Ile Glu Ala Phe Glu Pro Tyr Ile Asp Pro Ala
210 215

<210> 69
<211> 663
<212> DNA
<213> Unknown
5
<220>
<223> Metagenome - beta unit nitrile hydratase - M23dA12
10 <220>
<221> CDS
<222> (1) .. (663)
<223>

<400> 69
15 atg gac ggc gtg cac gac atg ggc ggc atg cac ggt ttc ggc aag gtc
48 Met Asp Gly Val His Asp Met Gly Gly Met His Gly Phe Gly Lys Val
1 5 10 15
0 gag ccg gaa gcg aac gag ccc gcc ttc cat gcg gaa tgg gaa ggc cgc
96 Glu Pro Glu Ala Asn Glu Pro Ala Phe His Ala Glu Trp Glu Gly Arg
20 25 30
25 tgc ctc gcg ctc aac cgc gcc atg ggt gcg atc ggc gcc tgg acc atc
144 Cys Leu Ala Leu Asn Arg Ala Met Gly Ala Ile Gly Ala Trp Thr Ile
35 40 45
30 gat gaa ggc cgt gcc ggc atc gag atc ctg ccg ccg gag att tat ctt
192 Asp Glu Gly Arg Ala Gly Ile Glu Ile Leu Pro Pro Glu Ile Tyr Leu
50 55 60
35 ggc agt tcg tac tat gga aaa tgg gcg ccg ccg ctg gag aat atg gtg
240 Gly Ser Ser Tyr Tyr Gly Lys Trp Ala Arg Arg Leu Glu Asn Met Val
65 70 75 80
0 gtc gca cgc ggg ttc gcg ggc gcc gat gaa ctc gcc gcg ggt cgc gca
288 Val Ala Arg Gly Phe Ala Gly Ala Asp Glu Leu Ala Ala Gly Arg Ala
85 90 95
45 gcg cgt ccc ggc aga tcg gtg aaa cga aag ctt acg gtc gcc gaa gtg
336 Ala Arg Pro Gly Arg Ser Val Lys Arg Lys Leu Thr Val Ala Glu Val
100 105 110
50 ccg cgc acg ctg acg cgc ggt tca ttt ttc cgc gag gca aca aag ccg
384 Pro Arg Thr Leu Thr Arg Gly Ser Phe Phe Arg Glu Ala Thr Lys Pro
115 120 125
55 gca cga ttt gcg gtc ggc gaa cgc gtg cgc acc agg aac att cat ccg
432 Ala Arg Phe Ala Val Gly Glu Arg Val Arg Thr Arg Asn Ile His Pro
130 135 140

480 gcg acg cac act cggttgcg cga tat gcg cgc ggc cat gtc ggc gtg
 480 Ala Thr His Thr Arg Leu Pro Arg Tyr Ala Arg Gly His Val Gly Val
 145 150 155 160
 5 atc gag gcg atc cgc ggt tgc cac gta ttt ccc gac tcg gtt gcg atc
 528 Ile Glu Ala Ile Arg Gly Cys His Val Phe Pro Asp Ser Val Ala Ile
 165 170 175
 10 ggc gcc ggc gag aac ccg caa tgg ctt tat acg gtg gtg ttc gaa ggc
 576 Gly Ala Gly Glu Asn Pro Gln Trp Leu Tyr Thr Val Val Phe Glu Gly
 180 185 190
 15 cgc acg ctg tgg ggc gat agc gcc gat ccg acg ctt aag gtc tcg atc
 624 Arg Thr Leu Trp Gly Asp Ser Ala Asp Pro Thr Leu Lys Val Ser Ile
 195 200 205
 0 gag gcg ttc gag ccg tat ctg gaa ccg gcc caa cca tga
 663 Glu Ala Phe Glu Pro Tyr Leu Glu Pro Ala Gln Pro
 210 215 220
 25 <210> 70
 <211> 220
 <212> PRT
 30 <213> Unknown
 <220>
 <223> Metagenome - beta unit nitrile hydratase - M23dA12
 35 <400> 70
 Met Asp Gly Val His Asp Met Gly Gly Met His Gly Phe Gly Lys Val
 1 5 10 15
 40 Glu Pro Glu Ala Asn Glu Pro Ala Phe His Ala Glu Trp Glu Gly Arg
 20 25 30
 45 Cys Leu Ala Leu Asn Arg Ala Met Gly Ala Ile Gly Ala Trp Thr Ile
 35 40 45
 50 Asp Glu Gly Arg Ala Gly Ile Glu Ile Leu Pro Pro Glu Ile Tyr Leu
 50 55 60
 55 Gly Ser Ser Tyr Tyr Gly Lys Trp Ala Arg Arg Leu Glu Asn Met Val
 65 70 75 80
 Val Ala Arg Gly Phe Ala Gly Ala Asp Glu Leu Ala Ala Gly Arg Ala
 85 90 95

Ala Arg Pro Gly Arg Ser Val Lys Arg Lys Leu Thr Val Ala Glu Val
 100 105 110

5 Pro Arg Thr Leu Thr Arg Gly Ser Phe Phe Arg Glu Ala Thr Lys Pro
 115 120 125

10 Ala Arg Phe Ala Val Gly Glu Arg Val Arg Thr Arg Asn Ile His Pro
 130 135 140

15 Ala Thr His Thr Arg Leu Pro Arg Tyr Ala Arg Gly His Val Gly Val
 145 150 155 160

20 Ile Glu Ala Ile Arg Gly Cys His Val Phe Pro Asp Ser Val Ala Ile
 165 170 175

25 Gly Ala Gly Glu Asn Pro Gln Trp Leu Tyr Thr Val Val Phe Glu Gly
 180 185 190

30 Glu Ala Phe Glu Pro Tyr Leu Glu Pro Ala Gln Pro
 210 215 220

35 <210> 71
 <211> 888
 <212> DNA
 <213> Unknown

40 <220>
 <223> Metagenome - beta unit nitrile hydratase - M49bD9

45 <220>
 <221> CDS
 <222> (1)..(888)
 <223>

50 <400> 71
 atg aac ggc gta cac gat ctt ggc ggg atg gat ggt ttc ggc cgg gtg
 48
 Met Asn Gly Val His Asp Leu Gly Gly Met Asp Gly Phe Gly Arg Val
 1 5 10 15

55 atg gcg gag gcg gac gag ccg gtc ttt cat gag ccc tgg gaa ggt cgc
 96
 Met Ala Glu Ala Asp Glu Pro Val Phe His Glu Pro Trp Glu Gly Arg
 20 25 30

144 gtc ttt gcg ctc aac atg ctc ggc atc ggg cgc gag ccc att ccg gtg

Val Phe Ala Leu Asn Met Leu Gly Ile Gly Arg Glu Pro Ile Pro Val
 35 40 45

5 gac gcg ctg cgc cat cgc att gag cgg atc gag ccg tgg cgc tat ctg
 192 Asp Ala Leu Arg His Arg Ile Glu Arg Ile Glu Pro Trp Arg Tyr Leu
 50 55 60

10 acg tcg agc tat tac gaa cga tgg ctg gcc gaa atg gag cag gcc atc
 240 Thr Ser Ser Tyr Tyr Glu Arg Trp Leu Ala Glu Met Glu Gln Ala Ile
 65 70 75 80

15 atc gat gcg ggc acg ctg act cct ggt gaa atc gat gcg cga atg ggc
 288 Ile Asp Ala Gly Thr Leu Thr Pro Gly Glu Ile Asp Ala Arg Met Gly
 85 90 95

20 gag ctc gaa acg gat cct gat cgc cca ctg cca agg act gat aac cct
 336 Glu Leu Glu Thr Asp Pro Asp Arg Pro Leu Pro Arg Thr Asp Asn Pro
 100 105 110

25 gag cat gcc gat ggg gtg gcg gcg ttg cgc gcc ggc agt ccc gta
 384 Glu His Ala Asp Gly Val Ala Ala Leu Arg Ala Gly Ser Pro Val
 115 120 125

30 acg cgc aag att cgc aag cag cgc ttc aca atc ggc gat cgg gtc
 432 Thr Arg Lys Ile Arg Lys Gln Pro Arg Phe Thr Ile Gly Asp Arg Val
 130 135 140

35 gta acg cgc aat ctt aat ccg cac ggc cat acg cgg ctg ccg cgc tat
 480 Val Thr Arg Asn Leu Asn Pro His Gly His Thr Arg Leu Pro Arg Tyr
 145 150 155 160

40 gcg cgc ggc aag cgc ggc gtc gta acg ctg cac cat ggc gca cat gtc
 528 Ala Arg Gly Lys Arg Gly Val Val Thr Leu His His Gly Ala His Val
 165 170 175

45 ttt ccg gat acg aac gcg cac ggg ctg ggc gag cat ccg cag cat ctc
 576 Phe Pro Asp Thr Asn Ala His Gly Leu Gly Glu His Pro Gln His Leu
 180 185 190

50 tat acg gtg cga ttt cct gcg cgc gag tta tgg agc gac gcg gcc gag
 624 Tyr Thr Val Arg Phe Pro Ala Arg Glu Leu Trp Ser Asp Ala Ala Glu
 195 200 205

55 ccg aaa gaa tcg ata atg atc gat ttg tgg gag agc tat ctt caa ccc
 672 Pro Lys Glu Ser Ile Met Ile Asp Leu Trp Glu Ser Tyr Leu Gln Pro
 210 215 220

gat atc ggc agc aaa gcg tcg tcg tcc gcg aaa ggc aaa gcg acg ccg
720 Asp Ile Gly Ser Lys Ala Ser Ser Ser Ala Lys Gly Lys Ala Thr Pro
225 230 235 240
5 aaa gtt aag ccc gca atg gcc aag gca acc gcc aag gta agc gtc tcg
768 Lys Val Lys Pro Ala Met Ala Lys Ala Thr Ala Lys Val Ser Val Ser
245 250 255
10 gcc aag gcc aaa act cgg gga aag gcg gcg ccg aag gag cgt cca aaa
816 Ala Lys Ala Lys Thr Arg Gly Lys Ala Ala Pro Lys Glu Arg Pro Lys
260 265 270
15 ctg aaa cct gcg cga gcg gcg acc tca gca gca tcc ggc ggc gaa aaa
864 Leu Lys Pro Ala Arg Ala Ala Thr Ser Ala Ala Ser Gly Gly Glu Lys
275 280 285
0 gct aag cga aag gcc aaa cga tga
888 Ala Lys Arg Lys Ala Lys Arg
290 295
25
<210> 72
<211> 295
<212> PRT
30 <213> Unknown

<220>
<223> Metagenome - beta unit nitrile hydratase - M49bD9

35 <400> 72
Met Asn Gly Val His Asp Leu Gly Gly Met Asp Gly Phe Gly Arg Val
1 5 10 15

0 Met Ala Glu Ala Asp Glu Pro Val Phe His Glu Pro Trp Glu Gly Arg
20 25 30

45 Val Phe Ala Leu Asn Met Leu Gly Ile Gly Arg Glu Pro Ile Pro Val
35 40 45

50 Asp Ala Leu Arg His Arg Ile Glu Arg Ile Glu Pro Trp Arg Tyr Leu
50 55 60

55 Thr Ser Ser Tyr Tyr Glu Arg Trp Leu Ala Glu Met Glu Gln Ala Ile
65 70 75 80
Ile Asp Ala Gly Thr Leu Thr Pro Gly Glu Ile Asp Ala Arg Met Gly
85 90 95

100 Glu Leu Glu Thr Asp Pro Asp Arg Pro Leu Pro Arg Thr Asp Asn Pro
105 110

5 Glu His Ala Asp Gly Val Ala Ala Ala Leu Arg Ala Gly Ser Pro Val
115 120 125

10 130 Thr Arg Lys Ile Arg Lys Gln Pro Arg Phe Thr Ile Gly Asp Arg Val
135 140

15 145 Val Thr Arg Asn Leu Asn Pro His Gly His Thr Arg Leu Pro Arg Tyr
150 155 160

20 165 Ala Arg Gly Lys Arg Gly Val Val Thr Leu His His Gly Ala His Val
170 175

25 180 Phe Pro Asp Thr Asn Ala His Gly Leu Gly Glu His Pro Gln His Leu
185 190

30 195 Tyr Thr Val Arg Phe Pro Ala Arg Glu Leu Trp Ser Asp Ala Ala Glu
200 205

35 210 Pro Lys Glu Ser Ile Met Ile Asp Leu Trp Glu Ser Tyr Leu Gln Pro
215 220

40 225 Asp Ile Gly Ser Lys Ala Ser Ser Ala Lys Gly Lys Ala Thr Pro
230 235 240

245 Lys Val Lys Pro Ala Met Ala Lys Ala Thr Ala Lys Val Ser Val Ser
250 255

45 260 Ala Lys Ala Lys Thr Arg Gly Lys Ala Ala Pro Lys Glu Arg Pro Lys
265 270

275 Leu Lys Pro Ala Arg Ala Ala Thr Ser Ala Ala Ser Gly Gly Glu Lys
280 285

50 290 Ala Lys Arg Lys Ala Lys Arg
295

55 <210> 73
<211> 630
<212> DNA
<213> Unknown

<220>

<223> Metagenome - beta unit nitrile hydratase - M6dE2

<220>

<221> CDS

5 <222> (1)...(630)

<223>

<400> 73

atg gac ggc att cat gat ctc ggt ggg atg agc ggg ttc ctc gtg
10 48
Met Asp Gly Ile His Asp Leu Gly Gly Met Ser Gly Phe Gly Leu Val
1 5 10 15

gag atc gag ccc gat gag ccg gtg ttc cac gag ccc tgg gag gcg ctg
15 96
Glu Ile Glu Pro Asp Glu Pro Val Phe His Glu Pro Trp Glu Ala Leu
20 25 30

gtt ttc gct ctg atg act ctc ggt atc ggg aag ctc ggc gcg tac aac
20 144
Val Phe Ala Leu Met Thr Leu Gly Ile Gly Lys Leu Gly Ala Tyr Asn
35 40 45

gcc gat gag tac cgc cac tcg atc gag ccg atg gat ccg gcc cac tac
25 192
Ala Asp Glu Tyr Arg His Ser Ile Glu Arg Met Asp Pro Ala His Tyr
50 55 60

ctt gcg gcg acg tac tac gag cgc atg ctc acc ggc gtc gca acg ctc
30 240
Leu Ala Ala Thr Tyr Tyr Glu Arg Met Leu Thr Gly Val Ala Thr Leu
65 70 75 80

ctc gtc gag aag aac gtc gtt gcc cgc gac gag ctc gaa gcg cgc gcc
35 288
Leu Val Glu Lys Asn Val Val Ala Arg Asp Glu Leu Glu Ala Arg Ala
85 90 95

ggc ggg ccc ttc ccg ctg tca cgg ccg gtg gcc gag ccg acg gcg
40 336
Gly Gly Pro Phe Pro Leu Ser Arg Pro Val Ala Glu Arg Pro Thr Ala
100 105 110

gac ctt cgg ccc cag cca caa cca cgc ttc gcg gtc ggg gat ccg gtc
45 384
Asp Leu Arg Pro Gln Pro Arg Phe Ala Val Gly Asp Arg Val
115 120 125

gtc gtg cgc gac atc cac ccg gcc ggg cat act cgt gtg ccg cgc tac
50 432
Val Val Arg Asp Ile His Pro Ala Gly His Thr Arg Val Pro Arg Tyr
130 135 140

gtg cgg ggc aag cgc ggg acc gtc gtg cac gtc gcg ccg aaa ttc tcg
55 480
Val Arg Gly Lys Arg Gly Thr Val Val His Val Ala Pro Lys Phe Ser
145 150 155 160

ttc ccc gac acg gcc gcg cac ggg ctg acc cat cg^g agc gag cac acg
 528
 Phe Pro Asp Thr Ala Ala His Gly Leu Thr His Arg Ser Glu His Thr
 165 170 175
 5
 tat cac gtg gaa ttc gtc gcg agt gac ctt tgg gcc gac gtg gcc ggg
 576
 Tyr His Val Glu Phe Val Ala Ser Asp Leu Trp Ala Asp Val Ala Gly
 180 185 190
 10
 agc aat gag agc gta ctc gtg gac ctg tgg gac ggc tat ctg gag ggc
 624
 Ser Asn Glu Ser Val Leu Val Asp Leu Trp Asp Gly Tyr Leu Glu Gly
 195 200 205
 15
 gca tga
 630
 Ala
 0
 <210> 74
 <211> 209
 <212> PRT
 25 <213> Unknown
 <220>
 <223> Metagenome - beta unit nitrile hydratase - M6dE2
 30 <400> 74
 Met Asp Gly Ile His Asp Leu Gly Gly Met Ser Gly Phe Gly Leu Val
 1 5 10 15
 35 Glu Ile Glu Pro Asp Glu Pro Val Phe His Glu Pro Trp Glu Ala Leu
 20 25 30
 40 Val Phe Ala Leu Met Thr Leu Gly Ile Gly Lys Leu Gly Ala Tyr Asn
 35 40 45
 45 Ala Asp Glu Tyr Arg His Ser Ile Glu Arg Met Asp Pro Ala His Tyr
 50 50 55 60
 Leu Ala Ala Thr Tyr Tyr Glu Arg Met Leu Thr Gly Val Ala Thr Leu
 65 70 75 80
 50 Leu Val Glu Lys Asn Val Val Ala Arg Asp Glu Leu Glu Ala Arg Ala
 85 90 95
 55 Gly Gly Pro Phe Pro Leu Ser Arg Pro Val Ala Glu Arg Pro Thr Ala
 100 105 110

Asp Leu Arg Pro Gln Pro Gln Pro Arg Phe Ala Val Gly Asp Arg Val
 115 120 125

5 Val Val Arg Asp Ile His Pro Ala Gly His Thr Arg Val Pro Arg Tyr
 130 135 140

10 Val Arg Gly Lys Arg Gly Thr Val Val His Val Ala Pro Lys Phe Ser
 145 150 155 160

15 Phe Pro Asp Thr Ala Ala His Gly Leu Thr His Arg Ser Glu His Thr
 165 170 175

Tyr His Val Glu Phe Val Ala Ser Asp Leu Trp Ala Asp Val Ala Gly
 180 185 190

20 Ser Asn Glu Ser Val Leu Val Asp Leu Trp Asp Gly Tyr Leu Glu Gly
 195 200 205

25 Ala

30 <210> 75

<211> 651

<212> DNA

<213> Unknown

35 <220>
 <223> Metagenome - beta unit nitrile hydratase - M25A18

40 <220>

<221> CDS

<222> (1)..(651)

<223>

45 <400> 75

atg cgc ggc acg cac gat ctc ggc gga ttg ccc gcc ggc ccg gtg gac

48

Met Arg Gly Thr His Asp Leu Gly Gly Leu Pro Ala Gly Pro Val Asp
 1 5 10 15

50 acc gct ccc cac gaa ccg acc ttc tgg gaa aag cag gtg gac gcg atc
 96

Thr Ala Pro His Glu Pro Thr Phe Trp Glu Lys Gln Val Asp Ala Ile
 20 25 30

55 cac ggc ctg ctc ggc gat tcc aag cgc cgc atc acg ctg cgc gac gag
 144

His Gly Leu Leu Gly Asp Ser Lys Arg Arg Ile Thr Leu Arg Asp Glu
 35 40 45

192 aac cgc ctc tat atc gaa tcg ctc ggc gac gac gtc tac aac acg ctc

	Asn Arg Leu Tyr Ile Glu Ser Leu Gly Asp Asp Val Tyr Asn Thr Leu			
	50	55	60	
5	ggc tat tac gag cgc tgg acc gcc gcc atg tgc cgc cag ctc atc gac			
	240			
	Gly Tyr Tyr Glu Arg Trp Thr Ala Ala Met Cys Arg Gln Leu Ile Asp			
	65	70	75	80
10	aag ggc gtg ctg acg cag gac gag atc gac gcc aag atc gcc gag ctg			
	288			
	Lys Gly Val Leu Thr Gln Asp Glu Ile Asp Ala Lys Ile Ala Glu Leu			
	85	90	95	
15	cgc gcc cgc ggc gtc ggc gcg gga cga cga aac ggc ctg cca acc			
	336			
	Arg Ala Arg Gly Val Gly Ala Gly Arg Arg Arg Asn Gly Leu Gln Thr			
	100	105	110	
20	gtg agc gcc gat ctg gcc gat ctg gcc atc gcg ccg cgc ttc gcc			
	384			
	Val Ser Ala Asp Leu Ala Ala Asp Leu Ala Ile Ala Pro Arg Phe Ala			
	115	120	125	
25	gcc ggc gac cgc gtg cgg gtg cgc gac gat tat ccg ccc ggg cac atc			
	432			
	Ala Gly Asp Arg Val Arg Val Arg Asp Asp Tyr Pro Pro Gly His Ile			
	130	135	140	
30	cgc acg ccg gtc tat gtg cgc ggc aag acg ggc gtg gtg acg cgc tgc			
	480			
	Arg Thr Pro Val Tyr Val Arg Gly Lys Thr Gly Val Val Thr Arg Cys			
	145	150	155	160
35	ttc ggc gcg ttc aag aac ccg gaa ttg ctc gcc atc ggc aag gac ggc			
	528			
	Phe Gly Ala Phe Lys Asn Pro Glu Leu Leu Ala Ile Gly Lys Asp Gly			
	165	170	175	
40	ctg ccc aag aaa att ctc tac gaa gtg cgc ttc aag cag acc gat ctc			
	576			
	Leu Pro Lys Lys Ile Leu Tyr Glu Val Arg Phe Lys Gln Thr Asp Leu			
	180	185	190	
45	tgg ccc gac tat gcc ggg ccg cgc acc gat acg ctg ctg atc gac atc			
	624			
	Trp Pro Asp Tyr Ala Gly Pro Ala Thr Asp Thr Leu Leu Ile Asp Ile			
	195	200	205	
50	tac gaa cat tgg ctg agc gac gcg tga			
	651			
	Tyr Glu His Trp Leu Ser Asp Ala			
	210	215		
55	<210> 76			
	<211> 216			
	<212> PRT			
	<213> Unknown			

<220>

<223> Metagenome - beta unit nitrile hydratase - M25A18

<400> 76

5

Met Arg Gly Thr His Asp Leu Gly Gly Leu Pro Ala Gly Pro Val Asp
1 5 10 1510 Thr Ala Pro His Glu Pro Thr Phe Trp Glu Lys Gln Val Asp Ala Ile
20 25 3015 His Gly Leu Leu Gly Asp Ser Lys Arg Arg Ile Thr Leu Arg Asp Glu
35 40 45Asn Arg Leu Tyr Ile Glu Ser Leu Gly Asp Asp Val Tyr Asn Thr Leu
50 55 600 Gly Tyr Tyr Glu Arg Trp Thr Ala Ala Met Cys Arg Gln Leu Ile Asp
65 70 75 8025 Lys Gly Val Leu Thr Gln Asp Glu Ile Asp Ala Lys Ile Ala Glu Leu
85 90 9530 Arg Ala Arg Gly Val Gly Ala Gly Arg Arg Arg Asn Gly Leu Gln Thr
100 105 11035 Val Ser Ala Asp Leu Ala Ala Asp Leu Ala Ile Ala Pro Arg Phe Ala
115 120 12540 Ala Gly Asp Arg Val Arg Val Arg Asp Asp Tyr Pro Pro Gly His Ile
130 135 14045 Arg Thr Pro Val Tyr Val Arg Gly Lys Thr Gly Val Val Thr Arg Cys
145 150 155 160
Phe Gly Ala Phe Lys Asn Pro Glu Leu Leu Ala Ile Gly Lys Asp Gly
165 170 17550 Leu Pro Lys Lys Ile Leu Tyr Glu Val Arg Phe Lys Gln Thr Asp Leu
180 185 19055 Trp Pro Asp Tyr Ala Gly Pro Ala Thr Asp Thr Leu Leu Ile Asp Ile
195 200 205Tyr Glu His Trp Leu Ser Asp Ala
210 215

5 <210> 77
<211> 657
<212> DNA
<213> Unknown

10 <220>
<223> Metagenome - beta unit nitrile hydratase - M50bD9

15 <220>
<221> CDS
<222> (1)..(657)
<223>

20 <400> 77
atg aac ggc atg cat gac atg ggc ggc atg cac ggc atg gga ccc att
48 Met Asn Gly Met His Asp Met Gly Gly Met His Gly Met Gly Pro Ile
10 1 5 10 15
0 cag atc gag aag gac gag tcg ccc ttc cat gcg cgc tgg gaa ggc cgg
96 Gln Ile Glu Lys Asp Glu Ser Pro Phe His Ala Arg Trp Glu Gly Arg
20 25 30
25 gcg caa gcg atg tac aac gcc att gcg gcc acg ggc aga ctg gtg ctt
144 Ala Gln Ala Met Tyr Asn Ala Ile Ala Ala Thr Gly Arg Leu Val Leu
30 35 40 45
30 ggc ggt aga ccc aca cgg gaa ggg ttc ccg ccg gcc gaa tat ctc cgc
192 Gly Gly Arg Pro Thr Arg Glu Gly Phe Pro Pro Ala Glu Tyr Leu Arg
50 55 60
35 atg agc tac tat gaa ttg ggt ttc agg gtg ctg gtc gag gac ttg gtc
240 Met Ser Tyr Tyr Glu Leu Gly Phe Arg Val Leu Val Glu Asp Leu Val
65 70 75 80
40 ctg aac ggt ttg gtg acg cgc gcg gaa atc acg agc ggc cgt ccg gca
288 Leu Asn Gly Leu Val Thr Arg Ala Glu Ile Thr Ser Gly Arg Pro Ala
85 90 95
45 aag ggg gct gca aag tcg acg ccc gca atc acc gcc acc gcg cag
336 Lys Gly Ala Ala Lys Ser Thr Pro Ala Ile Thr Ala Ala Thr Ala Gln
100 105 110
50 gca tat atg ttc gcg ctc aaa tcg acc ccg cga gac gta ccg gtc acg
384 Ala Tyr Met Phe Ala Leu Lys Ser Thr Arg Arg Asp Val Pro Val Thr
115 120 125
55 gcg cgt ttc caa gtc ggt cag cgt gtg cgc gcg cgc aac atc aat ccg
432 Ala Arg Phe Gln Val Gly Gln Arg Val Arg Ala Arg Asn Ile Asn Pro

	130	135	140
	gtc acc cat acg cgc ctg ccc cgt tac gcg cgc ggc aaa ttc ggc gtt		
480	Val Thr His Thr Arg Leu Pro Arg Tyr Ala Arg Gly Lys Phe Gly Val		
5	145	150	155
	atc gaa cgt gac cac ggt gtt tac agg ttc gac gat tcc ttt gcc acg		
528	Ile Glu Arg Asp His Gly Val Tyr Arg Phe Asp Asp Ser Phe Ala Thr		
10	165	170	175
	tcc ggc gac gag aag ccc cag cac gtt tat tct gtg cgc ttc gcg gcg		
576	Ser Gly Asp Glu Lys Pro Gln His Val Tyr Ser Val Arg Phe Ala Ala		
15	180	185	190
	cgc gaa cta tgg ggc gaa gcc gcg ccg cga gat gct gtc tat atc		
624	Arg Glu Leu Trp Gly Glu Ala Ala Pro Pro Arg Asp Ala Val Tyr Ile		
0	195	200	205
	gaa atc tgg gat gac aac ctt gag cca gcg tga		
657			
25	Glu Ile Trp Asp Asp Asn Leu Glu Pro Ala		
	210	215	
30	<210> 78		
	<211> 218		
	<212> PRT		
	<213> Unknown		
35	<220>		
	<223> Metagenome - beta unit nitrile hydratase - M50bD9		
	<400> 78		
40.	Met Asn Gly Met His Asp Met Gly Gly Met His Gly Met Gly Pro Ile		
	1	5	10
	15		
	Gln Ile Glu Lys Asp Glu Ser Pro Phe His Ala Arg Trp Glu Gly Arg		
	20	25	30
45			
	Ala Gln Ala Met Tyr Asn Ala Ile Ala Ala Thr Gly Arg Leu Val Leu		
	35	40	45
50	Gly Gly Arg Pro Thr Arg Glu Gly Phe Pro Pro Ala Glu Tyr Leu Arg		
	50	55	60
55	Met Ser Tyr Tyr Glu Leu Gly Phe Arg Val Leu Val Glu Asp Leu Val		
	65	70	75
	80		
	Leu Asn Gly Leu Val Thr Arg Ala Glu Ile Thr Ser Gly Arg Pro Ala		

cag cat caa gat ttg ccg tcc gat ctg gcg ctg cggtgtg aag tcg ctc
 144 Gln His Gln Asp Leu Pro Ser Asp Leu Ala Leu Arg Val Lys Ser Leu
 35 40 45
 5 gaa tcg ctg ctt gtc gag aag ggg ctg gtg gag cga gca gcg ctc gac
 192 Glu Ser Leu Leu Val Glu Lys Gly Leu Val Glu Arg Ala Ala Leu Asp
 50 55 60
 10 gcg ctg gtg gac acc tac gag cac aaa gtc ggg ccg cga aac gga gcg
 240 Ala Leu Val Asp Thr Tyr Glu His Lys Val Gly Pro Arg Asn Gly Ala
 65 70 75 80
 15 cgc gtt gtc gcg ccg gcc tgg gtt gac ccg gac tac aag caa cgg tta
 288 Arg Val Val Ala Arg Ala Trp Val Asp Pro Asp Tyr Lys Gln Arg Leu
 85 90 95
 20 ttc gcg aac ggt acc gcc gca gtc gcg gag ttc ggc tac tcc ggc tcg
 336 Phe Ala Asn Gly Thr Ala Ala Val Ala Glu Phe Gly Tyr Ser Gly Ser
 100 105 110
 25 cag ggc gct gac atc ccg gtc gtc gaa aac acg gcc act gta cat aac
 384 Gln Gly Ala Asp Ile Arg Val Val Glu Asn Thr Ala Thr Val His Asn
 115 120 125
 30 ctc gtc gtg tgc acg ctg tgc tct tgt tat ccc tgg ccg gtg ctg ggc
 432 Leu Val Val Cys Thr Leu Cys Ser Cys Tyr Pro Trp Pro Val Leu Gly
 130 135 140
 35 ttg ccg ccg gtc tgg tac aag tcc gcg ccc tat cgg tct cgc gtg gtg
 480 Leu Pro Pro Val Trp Tyr Lys Ser Ala Pro Tyr Arg Ser Arg Val Val
 145 150 155 160
 40 atc gat ccg cga ggt gtg ctg cgc gag ttc ggc gtg gtg ctg ccg gac
 528 Ile Asp Pro Arg Gly Val Leu Arg Glu Phe Gly Val Val Leu Pro Asp
 165 170 175
 45 cat atc gaa gtg cgt gtc tat gac agc acg gcg gag caa cgc tat cta
 576 His Ile Glu Val Arg Val Tyr Asp Ser Thr Ala Glu Gln Arg Tyr Leu
 180 185 190
 50 gtg ctg ccg gag ccg ccg gcc gga acc gaa aac ctg aca gaa gcg
 624 Val Leu Pro Glu Arg Pro Ala Gly Thr Glu Asn Leu Thr Glu Glu Ala
 195 200 205
 55 ctg gcg ctg ctg gtg acg cgc gac gcg atg att ggc gtg gcc aag gtc
 672 Leu Ala Leu Leu Val Thr Arg Asp Ala Met Ile Gly Val Ala Lys Val
 210 215 220

gcg ccg ccg gga ggc cgc gga tga
696
Ala Pro Pro Gly Gly Arg Gly
5 225 230

<210> 80
<211> 231
10 <212> PRT
<213> Unknown

<220>
<223> Metagenome - beta unit nitrile hydratase - M3aG10
15 <400> 80

Met Asp Pro Thr Arg Arg Ser Phe Leu Ala Ser Thr Val Ala Leu Thr
1 5 10 15

Gly Gly Ala Ala Ile Pro Asp Leu Ala His Ala Ala Asp His Asp His
20 25 30

Gln His Gln Asp Leu Pro Ser Asp Leu Ala Leu Arg Val Lys Ser Leu
25 30 35 40 45

Glu Ser Leu Leu Val Glu Lys Gly Leu Val Glu Arg Ala Ala Leu Asp
30 35 40 45 50 55 60

Ala Leu Val Asp Thr Tyr Glu His Lys Val Gly Pro Arg Asn Gly Ala
35 40 45 50 55 60 65 70 75 80

Arg Val Val Ala Arg Ala Trp Val Asp Pro Asp Tyr Lys Gln Arg Leu
40 45 50 55 60 65 70 75 80 85 90 95

Phe Ala Asn Gly Thr Ala Ala Val Ala Glu Phe Gly Tyr Ser Gly Ser
45 50 55 60 65 70 75 80 85 90 95 100 105 110

Gln Gly Ala Asp Ile Arg Val Val Glu Asn Thr Ala Thr Val His Asn
50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125

Leu Val Val Cys Thr Leu Cys Ser Cys Tyr Pro Trp Pro Val Leu Gly
55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140

Leu Pro Pro Val Trp Tyr Lys Ser Ala Pro Tyr Arg Ser Arg Val Val
60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160

Ile Asp Pro Arg Gly Val Leu Arg Glu Phe Gly Val Val Leu Pro Asp
65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175

His Ile Glu Val Arg Val Tyr Asp Ser Thr Ala Glu Gln Arg Tyr Leu
 180 185 190

5

Val Leu Pro Glu Arg Pro Ala Gly Thr Glu Asn Leu Thr Glu Glu Ala
 195 200 205

10

Leu Ala Leu Leu Val Thr Arg Asp Ala Met Ile Gly Val Ala Lys Val
 210 215 220

15

Ala Pro Pro Gly Gly Arg Gly
 225 230

0

<210> 81
 <211> 327
 <212> DNA
 <213> Unknown

25

<220>
 <223> Metagenome - p12K unit

30

<220>
 <221> CDS
 <222> (1)..(327)

<223>

35

<400> 81
 atg aaa gat agc ccg gtc ttt cgc gag ccg tgg gaa gcg cag gcg ttt
 48
 Met Lys Asp Ser Pro Val Phe Arg Glu Pro Trp Glu Ala Gln Ala Phe
 1 5 10 15

40

gac gaa gcg ttg gcg atc tcg ttg caa gac cgt ggc gtg ttc acg cga gac gaa
 96
 Ala Leu Ala Ile Ser Leu Gln Asp Arg Gly Val Phe Thr Arg Asp Glu
 20 25 30

45

tgg gcg gca ctc ggc gat gaa atc aag aag gcg caa gct gcc ggc
 144
 Trp Ala Ala Ala Leu Gly Asp Glu Ile Lys Lys Ala Gln Ala Ala Gly
 35 40 45

50

gat ccc gat acg ggc gag act tat tac cat cat tgg atg gca gcg ctc
 192
 Asp Pro Asp Thr Gly Glu Thr Tyr Tyr His His Trp Met Ala Ala Leu
 50 55 60

55

gaa cgg ctg att gca gcc aag ggt gtt gcc gat acg cag acg ctc gcg
 240
 Glu Arg Leu Ile Ala Ala Lys Gly Val Ala Asp Thr Gln Thr Leu Ala
 65 70 75 80

cgc aca cgc gac gcc tgg cag cac gcc tgt gcg cga acg ccg cat ggc
 288

Arg Thr Arg Asp Ala Trp Gln His Ala Cys Ala Arg Thr Pro His Gly
85 90 95

5 gcg cca atc gag cta aga ccg gac gac ttc agg aat tga
327 Ala Pro Ile Glu Leu Arg Pro Asp Asp Phe Arg Asn
100 105

10 <210> 82
<211> 108
<212> PRT
<213> Unknown

15 <220>
<223> Metagenome - p12K unit

<400> 82

0 Met Lys Asp Ser Pro Val Phe Arg Glu Pro Trp Glu Ala Gln Ala Phe
1 5 10 15

25 Ala Leu Ala Ile Ser Leu Gln Asp Arg Gly Val Phe Thr Arg Asp Glu
20 25 30

30 Trp Ala Ala Ala Leu Gly Asp Glu Ile Lys Lys Ala Gln Ala Ala Gly
35 40 45

Asp Pro Asp Thr Gly Glu Thr Tyr Tyr His His Trp Met Ala Ala Leu
50 55 60

35 Glu Arg Leu Ile Ala Ala Lys Gly Val Ala Asp Thr Gln Thr Leu Ala
65 70 75 80

40 Arg Thr Arg Asp Ala Trp Gln His Ala Cys Ala Arg Thr Pro His Gly
85 90 95

45 Ala Pro Ile Glu Leu Arg Pro Asp Asp Phe Arg Asn
100 105

50 <210> 83
<211> 321
<212> DNA
<213> Unknown

55 <220>
<223> Metagenome - p12K unit

<220>
<221> CDS
<222> (1)...(321)
<223>

<400> 83
 atg aga aca gtt gct gag caa atc gcg gct gat ctt gcg agt ccg gcg
 48
 5 Met Arg Thr Val Ala Glu Gln Ile Ala Ala Asp Leu Ala Ser Pro Ala
 1 5 10 15
 gcg att ccg cgc cgc aac ggc gag ccg gtc ttc gac gag cct tgg gaa
 96
 10 Ala Ile Pro Arg Arg Asn Gly Glu Pro Val Phe Asp Glu Pro Trp Glu
 20 25 30
 agt cgt gcg ttt ggg ata gcg gtc gcc ctt tcc gag ggt ggc ctc tat
 144
 15 Ser Arg Ala Phe Gly Ile Ala Val Ala Leu Ser Glu Gly Gly Leu Tyr
 35 40 45
 tca tgg gat gaa ttt cgc gat tgc ctg att gct gaa atc aca gcg gcg
 192
 0 Ser Trp Asp Glu Phe Arg Asp Cys Leu Ile Ala Glu Ile Thr Ala Ala
 50 55 60
 gat gcg cgc ggc gag cat acg agc tat tac gaa cgg ttt ctc gcc gcc
 240
 25 Asp Ala Arg Gly Glu His Thr Ser Tyr Tyr Glu Arg Phe Leu Ala Ala
 65 70 75 80
 ctg cag cat ctg ctc gcg gcc aaa cgc ctc tgc act ccc gat gaa gtc
 288
 30 Leu Gln His Leu Leu Ala Ala Lys Arg Leu Cys Thr Pro Asp Glu Val
 85 90 95
 gag cgg cgg atg aac act agc gca ggc acc tga
 321
 35 Glu Arg Arg Met Asn Thr Ser Ala Gly Thr
 100 105

 40 <210> 84
 <211> 106
 <212> PRT
 <213> Unknown

 45 <220>
 <223> Metagenome - p12K unit

 <400> 84
 Met Arg Thr Val Ala Glu Gln Ile Ala Ala Asp Leu Ala Ser Pro Ala
 50 1 5 10 15
 Ala Ile Pro Arg Arg Asn Gly Glu Pro Val Phe Asp Glu Pro Trp Glu
 20 25 30
 55 Ser Arg Ala Phe Gly Ile Ala Val Ala Leu Ser Glu Gly Gly Leu Tyr
 35 40 45

Ser Trp Asp Glu Phe Arg Asp Cys Leu Ile Ala Glu Ile Thr Ala Ala
50 55 60

5 Asp Ala Arg Gly Glu His Thr Ser Tyr Tyr Glu Arg Phe Leu Ala Ala
65 70 75 80

10 Leu Gln His Leu Leu Ala Ala Lys Arg Leu Cys Thr Pro Asp Glu Val
85 90 95

15 Glu Arg Arg Met Asn Thr Ser Ala Gly Thr
100 105

0 <210> 85
<211> 426
<212> DNA
<213> Unknown

25 <220>
<223> Metagenome - p12K unit

30 <220>
<221> CDS
<222> (1)..(426)
<223>

35 <400> 85
atg aca acc ttg agc cag cgt gaa gcg gcc ccc tcg gcc gag ctt ctt
48 Met Thr Thr Leu Ser Gln Arg Glu Ala Ala Pro Ser Ala Glu Leu Leu
1 5 10 15

40 gac cta ccg caa ctt cca agc gac acc gac ggc ccc gtc ttc gcg gaa
96 Asp Leu Pro Gln Leu Pro Ser Asp Thr Asp Gly Pro Val Phe Ala Glu
20 25 30

45 cct tgg gaa gcg gaa gcg ttt gcg ctt gcc gta agt ctt tca gag caa
144 Pro Trp Glu Ala Glu Ala Phe Ala Leu Ala Val Ser Leu Ser Glu Gln
35 40 45

50 gga cat ttc acg tgg aag gaa tgg gca gca acg ctc gcc gat gaa ctg
192 Gly His Phe Thr Trp Lys Glu Trp Ala Ala Thr Leu Ala Asp Glu Leu
50 55 60

55 gag ggc gcc aat cgc ggc gag ccg gat gac ggt acg cat tat tat
240 Glu Gly Ala Ala Asn Arg Gly Glu Pro Asp Asp Gly Thr His Tyr Tyr
65 70 75 80

288 gag tac tgg ctg acg gcc ctg gaa agg ctg gtt acg atc aag ggc ctg
Glu Tyr Trp Leu Thr Ala Leu Glu Arg Leu Val Thr Ile Lys Gly Leu

	85	90	95
	aca gat cag caa gcg atg cgc gag cgc aaa gag gcg tgg gaa gaa gcc		
	336		
5	Thr Asp Gln Gln Ala Met Arg Glu Arg Lys Glu Ala Trp Glu Glu Ala		
	100	105	110
	tat cgc cat acc ccg cat ggc gcg cca gtt gaa ctt atg tct ccg ctc		
	384		
10	Tyr Arg His Thr Pro His Gly Ala Pro Val Glu Leu Met Ser Pro Leu		
	115	120	125
	gat caa agc cgg ata gcc gaa gag gac agc gaa tcc tca tag		
	426		
15	Asp Gln Ser Arg Ile Ala Glu Glu Asp Ser Glu Ser Ser		
	130	135	140
	<210> 86		
	<211> 141		
	<212> PRT		
	<213> Unknown		
	<220>		
25	<223> Metagenome - p12K unit		
	<400> 86		
	Met Thr Thr Leu Ser Gln Arg Glu Ala Ala Pro Ser Ala Glu Leu Leu		
	1	5	10
			15
	Asp Leu Pro Gln Leu Pro Ser Asp Thr Asp Gly Pro Val Phe Ala Glu		
	20	25	30
35	Pro Trp Glu Ala Glu Ala Phe Ala Leu Ala Val Ser Leu Ser Glu Gln		
	35	40	45
	Gly His Phe Thr Trp Lys Glu Trp Ala Ala Thr Leu Ala Asp Glu Leu		
	50	55	60
	Glu Gly Ala Ala Asn Arg Gly Glu Pro Asp Asp Gly Thr His Tyr Tyr		
	65	70	75
			80
	Glu Tyr Trp Leu Thr Ala Leu Glu Arg Leu Val Thr Ile Lys Gly Leu		
	50	85	90
			95
	Thr Asp Gln Gln Ala Met Arg Glu Arg Lys Glu Ala Trp Glu Glu Ala		
	100	105	110
55	Tyr Arg His Thr Pro His Gly Ala Pro Val Glu Leu Met Ser Pro Leu		
	115	120	125

040067 AM

75

Asp Gln Ser Arg Ile Ala Glu Glu Asp Ser Glu Ser Ser
130 135 140

5